KU Leuven
Humanities and Social Sciences Group
Institute of Philosophy

Personal Identity and the Formal Self

Fauve LYBAERT

Supervisor:
Prof. A. Burms

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Philosophy

December 2013

TABLE OF CONTENTS

INTRODUCTION

'Who am I?' is one of the most basic, famous and reemerging questions in philosophy. It introduced many different quests – such as those that ask what we are made of, to which kind of being we belong, what is most essential to us, what is our purpose, how we ought to behave, and which changes in one would or would not cause one's death. I will focus on this latter question and examine what, in spite of all the psychological and physical changes we undergo, makes us identify ourselves as one numerically identical person throughout time. What makes me gather that I, now 27, was once a child organizing for safer traffic? Is there something in me that remained the same and makes me assume that I am still numerically the same entity, in spite of the differences between me, now, and this child, then?

This question echoes two major questions in the long-standing debate about the numerical identity of persons: (1) the metaphysical question 'what *constitutes* this numerical identity?' and (2) the epistemological question 'how do we *know* that we are numerically identical with a past or future being?' For John Locke – who, at the end of the seventeenth century, engendered an intense discussion on the constitution of diachronic personal identity with his *Essay Concerning Human Understanding* (Locke 1996 [1690]) – these two questions were inextricable. In a response to Christian philosophers, who assumed that what makes us who we are, must be our immaterial soul (which can survive our corporeal decay and death), Locke argued that this makes little sense, because these soul misses characteristics that could make us experience that it is present in someone, and so,

could give us the knowledge that this someone is the particular person to whom this soul belongs (Locke 1996, II.xxvii.14). Someone could say that he is Socrates, but we would have no means to verify this, just as I would not know if I am now still the same as I was a moment before. This is why Locke suggests another necessary and sufficient criterion for the preservation of the numerical identity of a person.

Locke defines a person as

> "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places" (Ibid, II.xxvii.9).

He continues that this person does this

> "only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it: it being impossible for anyone to perceive without perceiving, that he does perceive. When we see, hear, smell, taste, feel meditate, or will anything, we know that we do so…and by this everyone is to himself, that which he calls *self*" (Ibid).

Locke concludes that someone remains the same person as long as his consciousness, through which he is conscious of his sensations and thoughts at this moment, still makes him aware of what he experienced earlier:

> "as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person" (Ibid, II.xxvii.9).

The latter has often been interpreted as meaning that we have to remember what we did earlier, in order to really be the person who did this earlier. Locke provoked this interpretation. In his further discussion of personal identity, he emphasizes that persons are moral beings and that what constitutes persons, must therefore also

constitute them as moral beings. He then argues that our consciousness of what we did, aka our memory, is constitutive of persons as moral beings, because it allows them to feel that *they* did something, and, so, to feel responsible for it (Ibid, II.xxvii.20). In this passage Locke indeed identifies our consciousness of having done something with our memory, but this is not the case in the passage, quoted earlier (Ibid, II.xxvii.9) where he first identifies persons. There, Locke writes that for one to be a particular person, the consciousness with which one is conscious of one's own consciousness, now, must be the same as the consciousness with which one was conscious of one's earlier consciousness. This means that we must still be conscious of the earlier event. However, still being conscious of something is not necessarily the same as having a memory of something.

It remains unclear how we can determine whether the consciousness, through which we recall earlier events, is identical to the consciousness that made us conscious of these events when we first experienced them. Just as Locke asked what would make a soul into the same soul, we can ask Locke what would make a consciousness into the same consciousness. Yet, Locke was convinced that we feel or *know* when we still have the same consciousness as an earlier person, and he saw this as evidence in favour of his suggestion that a person *is* the same person as long as he has the same consciousness. Contrary to a mysterious soul, my consciousness allows me to feel that *I* feel something, and, so, that I am a *particular* self that can feel things. Consciousness is then both what makes us recognize ourselves as particular persons, and what constitutes us as these persons.

Locke's different insights concerning the person can be distinguished from one another. It is true that persons typically

consider themselves as one particular and diachronically existing entity. It is further likely that our capacity to consider ourselves as particular, diachronically existing entities, on the one hand, and the facts (1) that we are conscious of a great part of our consciousness and (2) that we have memories, on the other hand, are not independent of one another. However, this does not imply that these three forms of being aware of ourselves are identical. It is possible that we could not think of ourselves as being particular, diachronically existing entities, if we would not be aware of our consciousness, or would not have any memories, but this does not mean that our consciousness of being one diachronic entity is a consciousness of being consciousness, or, by extension, a consciousness of all that we have been conscious of. I will separately analyze in what our consciousness of being one diachronically identical entity consists, and what the elements are that allow for this consciousness.

The publication of Locke's *Essay Concerning Human Understanding* immediately generated further discussion. Part of the debate focused on the details of Locke's 'memory' criterion for personal identity. Joseph Butler argues that it would be wrong to say that the fact that a person still remembers his past proves that he is still the same, since we will only say that a person *remembers* something, when he actually did what he seems to remember. If not, we would not call it memory. According to Butler (1975 [1736], p. 100), remembering something presupposes that you were the person who did what you now apparently remember. Hence, it cannot first make it the case that you are this person. Thomas Reid agrees with Butler. He adds that even if memory can merely testify, and not constitute, personal identity, we would have to amend what Locke said about it. Reid judges that the

claim that we can only be a person, whose actions and thoughts we remember, is too strong. According to Reid, I can be one particular person, as long as there is a continuous chain of memories between me and that person. Reid (2010 [1785], Essay 3, chapter 6, pp. 147-148) offers the example of the general who can be said to have been a particular child who stole apples, even when he does not remember to have done so. For this, it suffices that the general remembers having been an officer, and that this officer remembers having stolen apples as a child.

Other philosophers – amongst whom Leibniz, Hume and Kant are most famous – did not so much focus on Locke's 'memory' criterion, as they developed their own thoughts on personal identity, in response to Locke and the Christian philosophers that preceded him. In his *New Essays on Human Understanding*, Leibniz (1981 [1704]) replied directly to each of Locke's paragraphs on personal identity. Like Locke, Leibniz looks at how we know that we have been a particular person. However, contrary to Locke, he does not judge that only the consciousness of our own consciousness gives us access to such knowledge. Leibniz observes that we sometimes know that we have done something, merely because others witnessed us doing this and told us about it (Leibniz 1981, II.xxvii.9). From this, Leibniz concludes that our body, as well, must have a role in what makes us stay the same person over time (Ibid). He concludes this because those who say that they know that we have done something, only *saw* that we did this; they have no immediate access to our consciousness of consciousness. They can only surmise that the person who did something then, and still has the same body, must be the same person.

Hume (1978 [1739]), in his turn, questioned if there is any such thing as a self with a diachronic existence, for which we can seek to determine the criteria of persistence. He notes that when he tries to observe this self through introspection, he finds no such entity. He is aware of certain sensations and thoughts, but he does not have an impression of one and the same self that has all these sensations and thoughts (Hume 1978, I.iv.6, p. 254). In an attempt to explain what makes us have the idea that we are one self, if not an impression of this self, Hume suggest that this may be caused by our observation of resemblances and the imagination of causal links between certain thoughts (Ibid, p. 260-261). Hume suggests that we are tempted to think that the reason for the occurrence of these resemblances and causal links would be given, if there is one and the same self that has these thoughts. So, our observation of sensations and thoughts could make us imagine that there is a persisting self. However, this self remains a fiction. We feign it. There is not one impression that corresponds with it.

Kant approached the question of numerical personal identity differently. In the third paralogism of his *Critique of Pure Reason*, Kant (1998 [1781], A361-A366) does not search for an observable numerically identical self, in order to determine whether it exists, what constitutes its identity, or what constitutes our knowledge of this identity. Instead, he admits that our thoughts and unity of consciousness refer to a subject which has them, but adds that this 'I' remains formal: we do not know anything about it, other than that it is the subject of these thoughts and unity of consciousness. Every thought has a thinker, and sometimes we hear and see something simultaneously, which presupposes that these perceptions are had by

the same perceiving subject. Yet, we cannot assume more about this subject. It is the subject of this thought, or united thoughts only, not a specific personality or other substance that remains the same. Herewith, Kant also criticizes Descartes (2004 [1639]), Locke's predecessor, who stated that, given that there is thinking, there must be a thinking thing, and immediately concluded from this that this thing must then be an immaterial substance constituting a particular person.

The debate above was revived during the second half of the twentieth century. Philosophers, mainly from the analytical tradition, started to fiercely debate over what constitutes the numerical identity of persons; whether our *idea* of what constitutes this identity corresponds with reality; what is special about self-consciousness, self-knowledge, and self-reference; and to what the 'I' in self-reference refers. Psychological reductionists, such as Derek Parfit (1984), return to Locke and Reid, in as far as they hold that the numerical identity of what we call persons is constituted by psychological continuity, made up of memory and other psychological links. At the same time, they return to Hume, by saying that a self does not truly exist. According to psychological reductionists, the self is nothing but a series of psychological relations.

The opposing camp of contemporary philosophers stresses the role of the body in the constitution of the numerical identity of persons. They do so for a variety of reasons. Some refer to the importance of our subjective point of view on the world – which is determined by the positioning of our body – for the constitution of our feeling of being a particular individual (Naomi Eilan 1995). Others argue that the body must be important for this constitution because it is needed for us to

be living human-beings (Olson 1997, Snowdon 1995) or because we could not be rational beings without it (Baker 2000). Philosophers like P.F. Strawson (1959), Saul Kripke (1972), Bernard Williams (1973b), David Wiggins (1976), and John McDowell (1997) attribute a more complicated role to the body, which comes closer to what Leibniz had in mind. They argue that, as a material and continuous entity, the body may not tell us anything about a person's personality, yet, it allows us to point to persons and to re-identify them in the spatio-temporal world in which we live.

A third group of contemporary philosophers focusses on a more transcendental self, such as Kant discussed, i.e. a self that is implied, but about whose characteristics we may not know much. Vendler (1984), who mainly tries to reinterpret Kant in this regard, is one of them. Wittgenstein (1958), Shoemaker (1963, 1968) and Evans (1982), who observe that we can correctly refer to ourselves, without thereby identifying ourselves through certain characteristics, can also be assigned to this group. The self about whom they write, is not necessarily the subject of a thought or unity of consciousness, as is the transcendental self of Kant. Yet, just like the self that Kant described in his third paralogism, it remains formal: we refer correctly to it, without knowing anything particular about it.

Rather than immediately determining what must metaphysically constitute the diachronic identity of persons, or how we, step by step, get to know whether we are identical to a person from the past, I will examine how we factually ascribe numerical identity to ourselves and others, and subsequently analyze what causes us to ascribe personal identity in this way. The factors that have a role in how we factually

ascribe personal identity, may be, but are not necessarily essential to its metaphysical constitution. We also do not need to consciously derive from them that we are one particular person. They may just function as background assumptions that allow us to ascribe personal identity as we do. After having determined how we ascribe numerical identity to persons, I will describe which kind of image of the self corresponds to the way in which we ascribe numerical identity to persons.

More specifically, in the first chapter, I will demonstrate how distinct philosophers assume that what makes us numerically identical must coincide with one, or more elements, which preserve our qualitative identity to the extent to which these qualities are important for us and characterize our particular personality. I argue that this assumption is incorrect: we sometimes judge that we are still numerically the same, even when we have changed substantially, qualitatively speaking. What constitutes the numerical identity of a person does not have to have a crucial, obvious role in the constitution of his typifying character.

After distinguishing the question of what constitutes the numerical identity of a person from what constitutes his qualitative identity, I explain, in the second chapter, why the psyche to which we typically attribute an important role in the constitution of the qualitative identity of persons cannot, on its own, constitute their numerical identity. I mention which characteristics of our body, considered as mere continuous material entity, render it more suitable to constitute this numerical identity. I also illustrate how we factually judge that persons remain numerically the same, as long as their one particular (continuous) body is capable of producing consciousness.

In the third chapter, I discern my approach of the role of the body in the constitution of personal identity from those of animalism and the constitution view. Contrary to protagonists of animalism and the constitution view my question is not what makes us stay the same throughout time ontologically speaking, but what makes us *consider* ourselves as the same diachronically existing being. Again, in opposition to defenders of animalism and the constitution view, I do not assume that, if the body has a role here, this must be because it is intrinsically important for a person's identity.

After establishing that we identify persons as numerically the same through one particular body that is still capable of producing consciousness, and after elaborating upon the specific role that the body could then have in this identification, I analyze which factors make us identify a person by his body. In the fourth chapter, I, in this regard, describe the role of material bodies and persons as primary particulars. In the fifth and sixth chapter, I show how the self-consciousness, which we often judge to be crucial for the preservation of personal identity is, in fact, informed by our idea that we, as persons, have just one particular physical history, which is that of one continuous material body. I show how this is the case for our episodic memory in the fifth chapter, and for our sensations and thoughts about ourselves as mental selves, in the sixth.

In the seventh and last chapter, I examine, to which kind of self we refer when we re-identify a person as numerically the same by his body. I argue that this self is a formal self. It is a self, to whom we refer, according to rules, and about whose character we do not need to know anything and cannot know anything definitive. Yet, while this self is

formal, it still determines how we think of ourselves as more substantial selves, i.e. as specific personalities.

Before developing my thesis, I wish to thank my supervisor Professor Dr. Arnold Burms for inspiring and advising me, throughout the years. I also benefitted from the advice of Professor Dr. Roland Breeur and Professor Dr. Nicolas de Warren, as well as from my fellow students at the KU Leuven, Boston College, Harvard University, M.I.T., and at international conferences. I further wish to acknowledge the financial support, which I received for this research project, from the Flanders Research Foundation and the Belgian American Educational Foundation. I am indebted to my home institution, The Higher Institute of Philosophy at the KU Leuven, for its excellent and authentic education, as well as the philosophy departments of Boston College, Harvard University, and M.I.T., which generously welcomed me as a Visiting Scholar. I am particularly grateful to Professor Dr. Jeffrey Bloechl and his family for their hospitality and invitation to the U.S.A., Dr. Cheryl Chen for acting as my supervisor during my time at Harvard, and Professor Dr. Galen Strawson, Jonathan Vogel, Alex Byrne, Susanna Siegel, Christine Korsgaard, Matthew Boyle, Sean Kelly, and Andreas Staiti for their helpful classes and advice. My thanks also goes to Vanessa Gorsuch for proofreading this dissertation. Further, I would like to thank Professor Dr. Russell Friedman, Martin Moors, and Bart Raymaekers for giving me the unique and motivating opportunity to teach philosophy. Last, but not least, I am deeply grateful to my partner, Sidar Ok; my friends, Sarit Larry, Véronique Euqinorèv, Gregory De Vleeschouwer, Andrej Zachar, Charlotte Lammertyn, Barbara Hoet,

and Anneleen Poll; my parents and family-in law, for their patience and support throughout my Ph.D. candidacy.

CHAPTER 1:
WHAT CONSTITUTES OUR NUMERICAL IDENTITY?

I. PERSONAL IDENTITY AS A NUMERICAL IDENTITY

To ask what constitutes a person's numerical identity, is to ask which elements have to stay in place for someone to remain numerically identical. These elements do not a priori have to guarantee his *qualitative* identity. Just as objects, persons may change qualitatively while still remaining numerically the same. Like a chair can remain numerically the same chair, even when it gets painted or loses a leg and thus changes qualities, a man can shrink with age and become more conservative, all the while we will say that this young and old man are one and the same person who has evolved over the years.

Still, philosophers who set out to discover the constitutive factors of our (idea of) numerical identity, have mostly come up with elements that make us qualitatively identical. In this chapter, I will sketch how this is the case for three prominent philosophers: John Locke, Derek Parfit and Bernard Williams. I will argue that their analysis is led by the thought that we must see the meaning and value for a specific personality of whatever could potentially constitute its personal identity. In the remainder of this dissertation, I will distance myself from this thought; separate the question of what constitutes our idea of numerical identity from that of what constitutes our idea of what

qualitative identity consists of; and examine whether elements which may, at first sight, not seem of importance to our personalities, could still constitute our idea of numerical personal identity, and why this would be the case.

1. John Locke

1.1 Locke defines the person

John Locke gave a major impulse to the philosophical discussion about what constitutes diachronic personal identity. Before answering this question, he considers it necessary to define what a person is. Locke defines the person as

> "a thinking, intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places" (Locke 1996, II.xxvii.9).

Locke continues that this being only does this

> "by that consciousness, which is inseparable from thinking…it being impossible for anyone to perceive, without perceiving, that he does perceive" (Ibid).

According to Locke, it is also this consciousness alone which constitutes our diachronic identity:

> "[n]othing but consciousness can unite remote existences into the same person" (Ibid, II.xxvii.23).

1.2 Locke's insight

A closer look at this claim reveals two interesting insights and a confusion of Locke's. Locke's first insight is that we first have to define what a person is, before we can find out what constitutes his numerical identity. Centuries later Geach (1967) explained why this is the case (without thereby referring to Locke). There *is* no such thing as an abstract or general numerical identity. Therefore we cannot tell what constitutes such mere numerical identity. What constitutes numerical identity, depends on of what we wish to determine the numerical identity. To determine what constitutes the numerical identity of a particular, we must therefore first know to which kind this particular belongs, as well as what makes it into a particular token within this kind (i.e. of what distinguishes it from other tokens of this kind). This needs to be clarified, so as to focus our attention. Clarifying this first does not lead to a circularity in the determination of what makes someone stay numerically the same. What makes something stay numerically the same may differ from what makes something belong to a certain kind, or what makes a token distinguish itself from the other tokens of that kind. It may, for example, be necessary to be born out of persons, so as to belong to the person kind; to have a separate body, so as to distinguish yourself from other tokens of the person kind; and to be psychologically continuous (say to only gradually and partially change in character and lose memories) so as to stay a numerically identical person. Still, you would first have to determine what a person is and so direct your attention to a particular kind of entity, before you could examine what constitutes his numerical identity.

Locke's second insight is less obvious and has been less accounted for in the contemporary debate on personal identity. It consists in his idea that persons are beings which can consider themselves as diachronically existing beings and that they can do so in a rather formal, or empty, way. Locke came to this statement by – just as he did for other ideas – seeking to both classify, as well as examine the origin of our idea of personal identity.

As we saw Locke classifies the person as

> "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places" (Locke 1996, II.xxvii.9).

Locke concludes that this idea must be given to us by our own consciousness of being conscious:

> "which [i.e. the considering of itself as itself] it does only by that consciousness, which is inseparable from thinking, and it seems to me essential to it: it being impossible for anyone to perceive without perceiving, that he does perceive" (Ibid).

What is interesting here, is that Locke initially refrains from making further claims about the nature of this self. Locke defines that a person must be conscious of its own diachronic being, but explicitly denies that this being must be a particular substance:

> "not being considered in this case, whether the same *self* be continued in the same, or diverse substances" (Ibid).

According to Locke, personal identity is merely constituted by the consciousness of consciousness and reaches as far

> "as this consciousness can be extended backwards to any past action or thought" (Ibid).

Again:

> "For it being the same consciousness that makes a man be himself to himself, *personal identity* depends on that only, whether it be annexed only to one individual substance, or can be continued in a succession of several substances. For as far as any intelligent being can repeat the *idea* of any past action with the same consciousness it had of it at first, and with the same consciousness it has of any present action; so far it is the same *personal self*. For it is by the consciousness it has of its present thoughts and actions, that it is self to it*self* now, and so will be the same *self* as far as the same consciousness can extend to actions past or to come…" (Locke 1996, II.xxvii.10).

And once more:

> "*Self* is that conscious thinking thing (whatever substance made up of whether spiritual, or material, simple, or compounded, it matters not) which is sensible, or conscious of pleasure and pain, capable of happiness or misery, and so is concerned for it*self*, as far as that consciousness extends" (Locke 1996, II.xxvii.17).

Herewith, Locke both follows and criticizes Descartes. Locke observes that we become aware of ourselves through our consciousness of being conscious, as Descartes remarked earlier:

> "Mais qu'est-ce donc que je suis? Une chose qui pense, qui doute, qui conçoit, qui affirme qui nie, qui veut, qui ne veut pas, qui imagine et qui sent" (Descartes 1956, 2nd med. §9).

> "il ne se peut pas faire que lorsque je vois, ou…lorsque je pense voir, que moi qui pense ne soit quelque chose" (Ibid, § 16).

Yet, as we saw, Locke does not conclude from this that this 'I' is an immaterial substance, contrary to Descartes:

> "…lorsque je considère mon esprit, c'est-à-dire moi-même en tant que je suis seulement une chose qui pense, je n'y puis distinguer aucunes parties, mais je me conçois comme une *chose seule, et entière*" (Descartes 1956, 6th med. § 33, my italics).

In his initial refusal to further specify what the self-conscious self is like, Locke is, in his turn, a predecessor of Kant. Just as Locke, Kant argues that there is no indication that this 'I' of self-consciousness is a substance. Kant went even further than Locke by stating that this 'I' which has a diachronic existence in our own consciousness or inner time, is hereby not granted such existence in an outer intersubjectively shared time:

> "…we cannot judge even from our own consciousness whether as soul we are persisting or not, because we ascribe to our identical Self only that of which we are conscious; and so we must necessarily judge that we are the very same in the whole of the time of which we are conscious. But from the standpoint of someone else we cannot declare this to be valid because, since in the soul we encounter no persisting appearance other than the representation 'I,' which accompanies and connects all of them, we can never make out whether this I (a mere thought) does not flow as well as all the other thoughts that are linked to one another through it" (Kant 1998, A 364).

While Locke admitted that the 'I' of which we become aware in our consciousness of being conscious is not one particular substance, he did hold that this consciousness of consciousness constitutes our diachronic identity, i.e. a person. According to Locke, distinct instances of thoughts are had by the same person as long as they are had by the same consciousness of consciousness:

> " 'tis plain consciousness, as far as ever it can be extended, should it be to ages past, unites existences, and actions, very remote in time, into the same person, as well as it does the existence and actions of the immediately preceding moment: so that whatever has the consciousness of present and past actions, is the same person to whom they both belong" (Locke 1996, II.xxvii.16).

As I will discuss later, it is not clear what should be in place for this consciousness to be the same. But the idea of such same consciousness

allowed Locke to say that an experience and a memory of this experience can be ascribed to the same person, given that the same consciousness of consciousness which made this person aware of the first experience, is what later makes him recall this experience. In contrast with this, Kant denies that there is an indication that the 'I', to which different instances of consciousness of consciousness refer, remains in some cases the same 'I' over (outer) time.

I agree with Kant that there is nothing in the 'I' to which every conscious act refers that makes it identical to an 'I' to which another conscious act refers. Yet, later in this thesis, I will return to the *formal* consciousness which we *do* have of being one *diachronic* self, and hereby acknowledge a seldom recognized insight of Locke. Without accepting a priori that this consciousness is solely constituted by our consciousness of consciousness as stated by Locke, I will examine if we have a formal idea of being a diachronic self and, if so, how this formal consciousness of being a diachronic self comes about, as well as what this says about our idea of being a diachronic self.

1.3 Locke's confusion: it is not because consciousness is typifying for persons that it alone constitutes numerical identity

With this in mind, we can return to Locke's application of his first insight to the case of persons. How does Locke define them as kind, token and numerically identical entity?

In defining persons as *kind*, Locke distinguishes them from man as a human organism, as well as from other substances which have to be continuous to remain. According to Locke, persons are selves which are, in this moment, conscious of themselves, through their awareness of being aware, and, at a later point in time, through their memory of what they experienced before. Following this view, persons can stop and resume their existence. We should not say that they are there during a deep, unconscious sleep, but we can say that they are the same persons as before the sleep, if they, after the sleep, remember what they did before (Locke 1996, II.xxvii.10). This is not the case for human organisms and other substances: their existence has to be continuous. Discontinuity means the end of the organism or substance.

According to Locke, persons and human organisms also differ from one another, in as far as a person's continued existence does not require that he lives through just *one* particular vital organism with a certain shape. A person could, in principal switch bodies. By contrast, a particular man or human organism cannot suddenly start living through the body of another man or non-human organism.

A last crucial difference, drawn by Locke, between human organisms and persons is that a human organism stays a human organism, even if its intelligence level drops beneath that of many animals, just as an animal stays an animal, no matter how intelligent it is; whereas a person only stays the same person as long as he thinks in a similar way as he used to do. A prince could start to live through the body of a cobbler and still be the prince, because he thinks in a way identical to how this prince used to think. But if a prince-shaped person no longer has the thoughts and way of behaving as he used to have, then he is no longer this prince-person (Ibid, II.xxvii.15).

Following Locke, persons are also moral beings. Their awareness of what they do, allows them to later recognize that they performed certain actions and take responsibility for these by accepting their consequences, be it reward or punishment (Ibid, II.xxvii.18).

In his analysis of persons as individual *token*-persons, Locke demonstrates how the same self-consciousness which is required to belong to the person-kind, also distinguishes persons from one another. He first excludes some other possibilities. The immaterial soul cannot individuate persons, because we do not know anything about this soul and it therefore does not provide us with a characteristic which could individuate a person (Ibid, II.xxvii.14). Bodies too fail to individuate persons. For, if we would notice that the person, now connected with the body of a former cobbler, has the behaviour and memories of a former prince and vice versa, then we would no longer identify the prince and the cobbler by their bodies. Instead, we would say that the prince and the cobbler must have changed bodies (Ibid, II.xxvii.15). Locke considers self-consciousness to be a better candidate for the individuation of persons. It allows persons to experience that they differ from one another: they experience their own experiences as theirs and because they have the immediate consciousness that their experiences are theirs, they also remember them as theirs.

So far, Locke's analysis of a person as kind and token is particular, but acceptable. The problem in his analysis, with which I wish to deal here, lays in his quick transition from his idea of what a person is, as kind and as token, to his conclusion about what constitutes the numerical identity of persons. In the first substantial paragraph on personal identity in his *Essay Concerning Human Understanding* (Locke

1996, II.xxvii.9), this transition is immediate and without further justification. From the thought that a person as kind is "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places," and the idea that a person as token individuates itself from others by "perceiving…that he does perceive," Locke immediately concludes that this same consciousness which typifies him as kind, and individuates him as token, must also constitute him as diachronically existing, numerically identical entity: "as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that *person*; it is the same *self* now it was then." Here, Locke assumes that, what constitutes the numerical identity of a person, must be of clear importance for the qualitative character of this person as kind and as token.

I disagree. It is not because consciousness is a very important quality of persons, as kind and as token, that this same consciousness constitutes the person's numerical identity. Say that we agree that diachronic self-consciousness is both necessary for someone to belong to the kind 'person' and for a person to be distinguished from other persons. Say that we also agree that distinct experiences can be assembled as mine by my coconsciousness – I experience them as mine and so remember them as mine. Even then, it would be too quick to conclude, as we saw Locke do above, that this coconsciousness alone constitutes diachronic identity. Other possibilities are that

(1) coconsciousness is not a sufficient condition for the constitution of diachronic self-consciousness – coconsciousness itself might be constituted by, to coconsciousness external, reference points, or it may

have to make us aware of an event at which we were physically present, to count as real coconsciousness;

(2) coconsciousness is not a necessary condition for diachronic self-consciousness – the latter may for instance be constituted by our reliance on testimony rather than on information coming 'from the inside'[1];

(3) our numerical identity is not constituted by our most important quality – perhaps consciousness is what we value most in persons, yet another of their properties which we do not value as such constitutes their numerical identity.

I will elaborate on these possibilities in the subsequent chapters of this dissertation. Here, my aim was merely to point out how Locke ignores these possibilities – a fact which seems to be caused by his valuation of a particular quality in persons. In an answer to his historical predecessor Descartes, who distinguished divisible, extended matter from indivisible, unextended thinking, and to the subsequent Christian thinkers, who assumed that the substance of this thinking must be an immaterial soul, Locke singled out consciousness as the most typifying characteristic of a person, which can unify him even if he is not a substance. Locke's focus on the typifying character of this consciousness, as well as on the apparent capacity of our consciousness to unify ourselves, makes him look no further for constitutive elements of our numerical identity.

Another proof of Locke's association of what constitutes personal identity with what we find important in persons is his quick identification of our diachronic self-consciousness that he first

[1] Cfr. Leibniz' critique of Locke in Leibniz 1981, 236 §9.

described in rather formal terms, with a consciousness of a very particular personal history, whose existence justifies our moral and juridical judgments. While Locke first describes a person as "a…being that…can consider itself as itself" (Locke 1996, II.xxvii.9), and does not define further in what this self-consciousness consists, except that it is a consciousness of being "the same thinking thing in different times and places" (Ibid), he later means to demonstrate that our moral and juridical judgments are justified because they take into account that we are only the same self as long as we are conscious of this self.[2] This latter consciousness is a consciousness of having been a particular personality, who had specific thoughts and performed specific actions. It is no longer a rather formal awareness of being the same self, without knowing more about this self. Instead, this consciousness consists in a set of specific memories of having a particular history. We only punish people for those actions that they remember doing. This justifies our punishment. For, if we would punish people for something they do not remember doing, it would be to them as if we punished them for something someone else did.[3] Locke makes it appear as if his analysis of what constitutes the numerical identity of persons comes first and the fact that our moral behaviour is justified because it takes this constitution into account only second.[4] However, this is hard to

[2] "In this *personal identity* is founded all the right and justice of reward and punishment; happiness and misery, being that, for which everyone is concerned for *himself*…" (Locke 1996, II.xxvii.18).

[3] "…to punish *Socrates* waking, for what sleeping *Socrates* thought, and waking *Socrates* was never conscious of, would be no more of right, than to punish one twin for what his brother-twin did, whereof he knew nothing" (Locke 1996, II.xxvii.19).

[4] "In this *personal identity* is founded all the right and justice of reward and punishment; happiness and misery, being that, for which everyone is concerned for *himself*…" (Locke 1996, II.xxvii.18, my underlining).

maintain. First, Locke mentions that the term person is "a forensic term appropriating actions and their merit; … [which] belongs only to intelligent agents capable of a law, and happiness and misery" (Locke 1996, II.xxvii.26). Further, Locke only starts describing the consciousness of consciousness, which is necessary to be a person, as the *memory* of having *a particular history*, when he considers the person as moral entity who has to be able to ascribe specific actions to himself.[5] From this we can derive that Locke's interpretation of the, for personal identity, constitutive consciousness of consciousness, as a form of a memory, is influenced by his idea that we are moral entities. In the end, Locke's thought of what constitutes personal identity is influenced by his idea that persons are accountable beings.

I still agree with Michael Ayers (1996, pp. 266-268), a major interpreter of Locke, that Locke contends that personal identity is constituted by an ontological relation, i.e. coconsciousness, and not just by some kind of moral and legal bound that makes us take responsibility for 'ourselves', just as such bound can make us take responsibility for things which do not belong to our person, say our children or house. But the elements which Ayers cites to demonstrate that this relation is first and foremost ontological, I take to illustrate how what Locke proposes to be the ontological relation that

[5] Locke first interprets this consciousness of consciousness as memory, eleven paragraphs after he first mentioned its role in the constitution, and only two paragraphs after he describes the person as a moral being, in the middle of further considerations about this being as moral: "suppose I wholly lose the memory of some parts of my life, beyond a possibility of retrieving them, so that perhaps I shall never be conscious of them again; yet am I not the same person…? To which I answer, that we must here take notice what the word *I* is applied to, which in this case is the man only….human laws not punishing the *madman* for the *sober man's* actions, nor the *sober man* for what the *madman* did, thereby making them two persons…" (Locke 1996, II.xxvii.20, my underlining).

constitutes personal identity is influenced by his idea that we are moral beings. Ayers mentions how Locke's "explicit identification of consciousness of the past with memory or recollection …[and] his quasi-Cartesian explanation of 'consciousness' in terms of the necessary self-consciousness of mental operations" show that "[i]t is memory in a straight-forward sense which for Locke both ties past actions to the present" (Ayers 1996, p. 267). This memory is supposed to excite "such moral concern as the acknowledgement of guilt" (Ibid) and show that "[c]onsciousness is distinct from conscience, and prior to it" (Ibid). I argue that Locke identifies coconsciousness with memory because he assumes that the ontological relation which constitutes a person, must also constitute him as a moral entity. If this were not the case, Locke could have interpreted consciousness differently.

We have then found one other indication of the fact that Locke assumes that the numerical identity of persons must be constituted by an element which also constitutes him as a particular qualitative being. A last evidence in this regard, is Locke's quick switch from the term 'personal identity,' which he had been using, and refers to the numerical identity of persons, to the term 'personality', which refers to the person as a particular character, in the following quote: "punishment is annexed to personality, and personality to consciousness" (Locke 1996, II.xxvii.22). Herewith Locke wishes to demonstrate that the ontological element which constitutes personal identity, also constitutes us as moral entities. But by using the term 'personality' instead of 'personal identity' Locke does not so much illustrate that we are rational in making sure that the person we punish is ontologically the same as the person who committed the personal

acts, as he conveys his own assumption that an ontological numerical identical person is a particular personality who is a particular qualitative and moral being, in as far as he has done and thought specific things, and is still conscious of these, or still remembers these.

In the subsequent paragraphs, I turn to the contemporary philosophical debate on personal identity. Just like Locke, the two first protagonists and opposing parties in this debate, Derek Parfit and Bernard Williams, offer fine and new insights on the matter, yet still end up confounding criteria for the numerical identity of persons with those for their qualitative identity. This covers up their first insights. I will highlight what Parfit and Williams say in this regard. This should further clarify what it means to search for (our) criteria for the numerical identity of persons. It will also give a short overview of the two major positions in the contemporary debate on personal identity.

2. Derek Parfit

2.1 "Identity is not what matters"

Derek Parfit's thought on personal identity can be captured in one of his own phrases: "identity is not what matters" (Parfit 1984, pp. 215-217, 224-225, 241, 245, 255, 273-275; 1995, pp. 29, 33-34, 44; 2007 passim). He attaches both an ontological and a moral meaning to this expression.

The ontological claim is that a particular person is not one determinate being who is either there or not there. A person is not at all times fully identical to its past selves. Instead, it is connected to them to a greater or lesser degree. According to Parfit, identity is thus not crucial in the metaphysical make-up of persons. In this matter, it matters not.

With regard to our attitudes or moral, Parfit judges that what we really value (or ought to value as rational beings) in our personal survival is in fact not that *we* survive, i.e. that our identity is preserved. Even in our apparent valuation of the continuation of our personal identity, identity is not what matters.

I will elaborate on Parfit's ontological claim first and return to his moral claim in the subsequent paragraph.

2.1.1 The metaphysical constitution of persons: identity is not what matters

With regard to personal identity, Derek Parfit is a constitutive reductionist (Parfit 1995, p. 16; 1999). Constitutive reductionists are metaphysical reductionists, not conceptual reductionists. They claim that persons are not separately existing entities over and above their physical and/or psychological continuity, even though we may not be able to get rid of the *term* 'person' when we want to give a complete description of the world.

Another way to state what constitutive reductionists hold is the following. They claim that, what makes different experiences belong to one person, is not the fact that they belong to a single separately

existing entity. Rather, what makes experiences intrapersonal should be explained in terms of other facts, such as the fact that they are psychologically continuous with one another or the fact that they are associated with a single body.

A metaphysical non-reductionist, on the other hand, claims that persons are separately existing entities over and above their physical and psychological continuity. An example of a metaphysical non-reductionist would be someone who identifies persons in accordance with their soul and does not take this soul to be fully constituted by any combination of further entities.

With regard to personal identity there are two major kinds of constitutive reductionists. Some reductionists, like Bernard Williams (1973b), argue that a person stays the same person as long as there is a certain degree of physical continuity. Other reductionists, like Derek Parfit, hold that a person stays the same as long as there is a certain degree of non-branched psychological continuity. Parfit (1984, p. 216) claims specifically that

> "[w]e are not separately existing entities, apart from our brains and bodies, and various interrelated physical and mental events. Our existence just involves the existence of our brains and bodies, and the doing of our deeds, and the thinking of our thoughts, and the occurrence of certain other physical and mental events. Our identity over time just involves (a) relation R – psychological connectedness and/or psychological continuity – with the right kind of cause, provided (b) that this relation does not take a 'branching' form, holding between one person and two different future people."

With "our existence," Parfit (1984, p. 202) means our factual existence, not what is principally necessary for a person's existence. According to Parfit, we factually live through one body, but could in principle continue to live through another body. With our "identity

over time," he means "what this identity necessarily involves, or consists in" (Ibid). "Psychological connectedness" consists of direct psychological links (Ibid, p. 206), such as those between an intention and the acting on this intention, or an experience and the memory of this experience. "Psychological continuity is the holding of overlapping chains of strong connectedness" (Ibid). Psychological connectedness is a matter of degree. There can be more or less direct psychological connections between different person-stages (i.e. the same person at different moments). According to Parfit, this degree has to be sufficient for us to still speak of the same person (Ibid). However, a person can still be called numerically identical to a person at an earlier stage, when there is no direct psychological link between these two person-stages. This is the case, when he is psychological continuous with this person or, in other words, when there are enough overlapping chains of psychological connectedness between the different person-stages that lead back to the first person-stage of this person. "The right kind of cause" would for some be a normal cause (Ibid, p. 207), such as a natural character change during the teenage years of a person, as opposed to an artificially produced character change, say by tampering with the brain. For Parfit, any cause will do as the right kind of cause (Ibid, p. 208). The condition "that this relation does not take a 'branching' form" is there to allow for the numerical identity of a person. One person can never be numerically identical to two distinct persons. If A is numerically identical with B and C, then B and C would have to be numerically identical with one another. If they are two different branches of a basic stream of psychological continuity, this is impossible: they will each develop another history and so not be numerically identical.

It is worth a note that Parfit's idea of what constitutes psychological unity – i.e. a series of sufficient psychological connections between specific thoughts, experiences and behaviour – is specific and not the only way to conceive of this unity. Parfit sees the unity of consciousness as a grand total of particular thoughts, experience and behaviour which are united by memory or meaningful psychological links such as those between intention and action. This does not yet explain how we integrate what we hear, see, smell, etc. – another instance of unity of consciousness still discussed by phenomenologists and analytic philosophers.[6] Neither does it say anything about the more formal form of self-consciousness, mentioned above, which we also have of being one diachronically existing entity, and which we cannot just assume to be constituted by an awareness of our specific psychological history. In subsequent chapters, I will return to how the unity of consciousness, which Parfit describes, and the formal self-consciousness relate to one another, and what this can mean for how our idea of personal identity is constituted. Here, I constrain myself to explaining how the nature of psychological unity, as conceived by

[6] About these phenomena, Parfit says that they belong to one state of consciousness, which implies that they belong to one subject, but not that they belong to one person (Parfit 1984, pp. 248-252). According to Parfit, one cannot maintain that all these experiences are united, as long as one person can say 'they are mine'. This would ignore the fact that persons sometimes have a divided mind. These persons have different streams of consciousness, which are in themselves united, but through which they are not aware of the other stream. Explaining the unity in these respective streams of consciousness, by saying that they are both mine, would make them into one, and ignore the possibility that "in having each of these two sets of experiences, I am unaware of having the other" (Ibid, p. 249). I will return to the possibility that one person has different streams of consciousness. It is real and, for instance, the case for epilepsy patients, whose right and left hemisphere are operationally split, to minimize to impact of their seizures. These patients can become aware of something through their right hemisphere, of which the consciousness produced by their left hemisphere remains unconscious.

Parfit, proves – again according to Parfit – that identity is not what matters in the ontological analysis of persons.

A first evidence for Parfit, in this regard, is the fact that psychological connectedness, which suffices to constitute personal 'identity,' is a matter of degree. Parfit reasons as follows. When there has to be a sufficient degree of psychological connectedness between person-stages so as to attribute them to numerically the same person, but there is no precise number which indicates how many connections constitute a sufficient degree, we may in some cases not know whether certain person-stages belong to the same person or not. Parfit (Ibid, pp. 213-214, 233) says not to worry about these borderline cases. While we in these cases cannot decide whether two person-stages are to be attributed to the same person or not, we still know everything there is to know about these cases. We know which ontological elements and relations we have in front of us. Given that identity is never a further fact than a sufficient degree of these relations, knowing that identity is involved, would not add anything to our ontological knowledge of the world. In our ontological analysis of this case identity is not what matters. Moreover, if we for one reason or another choose to decide whether or not numerical identity is still involved in these borderline cases, then identity would still not matter. The choice would be arbitrary (Ibid, p. 241). An identity arbitrarily attributed is not an identity that matters.

2.1.2 The unimportance of personal identity: identity is not what matters in our concern to survive

Parfit's phrase 'identity is not what matters' does not only say something about his view on the *nature* of personal identity (which proves to be reducible to a set of other relations). It also summarizes his view on the *importance* of personal identity. Parfit (1984, pp. 261-263, 272, 279, 309; 1995, p. 15; 2007) denies that the continuation of our unique personal existence is what really matters in our concern for survival. He argues that, when we think about it rationally, we will see that all we should really care about when we care about survival (be it our own or that of someone else), is the continuation of psychological continuity (1984, pp. 215, 216, 262, 263, 272, 292, 313; 1995; 2007). Parfit reasons that the psychological characteristics of someone provoke our love for and attachment to him in the first place. If the memory of this person is maintained and the character that typifies him only changes gradually, then we can expect that he will continue to pursue the projects that are important to him and to us.[7]

[7] More than twenty years after having published *Reasons and Persons*, Parfit (2007) admitted that this reasoning was too hasty and that he should distinguish our desire to merely survive from our desire to survive in what we consider to be a good way. He repeats that we can wish for a survival, in which psychological continuity is preserved because of the reasons here mentioned, but he admits that we can also have an interest in survival *tout court*. According to Parfit (2007) we want others to survive because of their specific qualities, while our desire to personally survive is not necessarily based on, or reducible to, a wish that some of our projects or qualities do not go extinct. We can just not want to die. This is an interesting line of thought, which I will explore further in this dissertation. However, here it only deserves a footnote, because it does not profoundly change Parfit's thoughts about what matters in survival. He continues to defend that personal identity is not what matters in our survival because persons can be reduced to a set of other relations.

There are several reasons why this psychological continuity does not imply identity. The first is that the unity of consciousness, which is constituted by this psychological continuity, does not stay identical, but can be present to a greater or lesser degree (1984, pp. 231, 298-299; 1995). There can be more or fewer memories, bigger or smaller changes in character or knowledge, more or less acting on previous intentions, etc.. Secondly, Parfit contends that this psychological continuity could, in principle, branch, split like an amoeba, or be downloaded from the current brain and nervous system and then again be uploaded in one or more other bodies (1984, pp. 245-248, 254-260, 262, 266, 269-270, 299-302). In each of these scenario's, distinct concurrent beings, which find themselves at different places and are psychological continuous with one previous being, would come into existence. In spite of their continuity with one previous being, these beings would not be numerically identical with one another.

2.1.3 Parfit's claims regarding the nature and importance of personal identity are different, but related

Parfit (D Parfit April 8th 2010, personal communication) insists that his claims about the nature of personal identity should be distinguished from his claims about its importance. The former view is supposed to merely follow from his reductionism and argument for reductionism and in no way to be determined by his view of what matters in personal identity.

However, Parfit (1984, p. 216) explains that if someone accepts his claims about the nature of personal identity, rational thinking requires

him to also accept his claims about the importance of personal identity: if we do not hold that persons are separately existing entities, but instead that they can be reduced to some physical matter and psychological continuities which continue in degrees, then we should also give up the idea that identity is what is important in our survival (Ibid).

2.2 Parfit's insight

Even though Parfit argues that psychological continuity really is, and should be, what we value when we seem to value identity, he does not make the same mistake as Locke. While Parfit recognizes that we consider psychological states and continuity to be important for the beings whom we call persons (1984, pp. 215, 216, 262, 263, 272, 292, 313), he denies – in contrast with Locke – that these states and continuity also constitute the numerical identity of persons. He even explains why we are, in fact, mistaken when we think to here be concerned with a certain *identity* (cfr. the matter of degree and branching). Herewith, Parfit clearly distinguishes the question of the qualitative and numerical identity of persons.[8]

[8] Parfit also explicitly draws this distinction. See, for example, Parfit 1984, pp. 201-202, Parfit 1995, pp. 13-14 and Parfit 2012.

2.3 Parfit's confusion

However, Parfit only distinguishes the questions of numerical and qualitative identity at that very moment, i.e. at the moment where he identifies the qualities that we value in a person and explains why these qualities are no guarantee for the numerical identity of a person. He confounds them in the remainder of his discussion of personal identity where he assumes that if something is to constitute personal identity, this must be a quality that is of clear importance and value to the person. This becomes clear when he does not truly consider an alternative for psychological continuity as a potential constituent of the person.

This is, for instance, illustrated by his discussion of a human being who changes gradually, psychologically speaking (Parfit 1984, pp. 231-233). Parfit claims that this being could no longer be the same person at the end of the spectrum of psychological changes. Here, he assumes that it is the psyche, clearly important for a person and his qualitative identity that also constitutes him as a numerically identical person. With regard to the same case, Parfit says that a person cannot become a different person, if only some of his memories fade, or characteristics change (Ibid). Hereby, he assumes that such small changes are not important enough to a person, so as to change his identity.

Parfit further confounds the questions of numerical and qualitative identity, when the fact that the set of psychological relations that constitutes the person is not determinate, makes him conclude that personal identity is not determinate (Parfit 1984, p. 279; 1995, pp. 21-22). This is also the case when he derives, from the same fact, that we

should not care about identity when we care about survival (Parfit 1984, pp. 261-263, 272). Parfit does not consider that another element, which is not (or less) important for the qualitative identity of a person, could constitute personal identity and make it determinate.

Parfit does shortly examine whether the body could help constitute the person, but he only seriously considers the potential role of the brain (Parfit 1984, pp. 204, 253-255, 269-271, 284) in as far as it helps constitute the psyche, which characterizes the person as type and token.

In all these cases, Parfit assumes that only qualities that we value in persons can constitute personal identity and make it determinate. He concludes that if the psychological continuity which we value in persons cannot constitute numerical identity, then personal *identity* cannot be what matters – be it ontologically, or to us personally. He does not consider that something that we at first sight do not value that highly in a person could potentially constitute his numerical identity. By not taking this possibility into account, Parfit in fact lets his view of the nature of personal identity be determined by what he takes to be important for our survival.

By concluding that personal *identity*, as such, cannot matter truly, Parfit fails to account for the fact that we can wish to survive, even when we do not know whether we will continue to have any of the qualities which we currently value in ourselves.[9] So, he for instance fails to account for the existing wish of some persons never to have euthanasia done to them, regardless of in which living state they may

[9] Parfit (2007) actually discovers this later, but does not revise his metaphysical positions because of it. I think that he should.

one day find themselves. These persons are concerned about the potential future being who would be numerically identical to them, not about a being who would have the same qualities as those they value now in themselves.

In the following chapter, I will examine whether elements, other than the ones we value most in persons, could help constitute our idea of their numerical identity and, if so, which elements these could be. Here, I foremost wish to point out that there is one sense of 'identity is not what matters' for which Parfit does not account. There is a possibility that our idea of personal identity is constituted by something that we do not value and therefore not by what matters (most) to us. Given the fact that Parfit explicitly mentions different interpretations of the idea that identity is not what matters to us, it is quite surprising that he does not consider this one. Just as Locke, Parfit immediately assumes that we are only prepared to say that a person stays numerically the same, as long as he does not lose a specific set of those qualities that we value most in a person. This makes Parfit not *truly* consider that the elements constitutive of the qualitative and numerical identity of persons may differ.

3. Bernard Williams

Bernard Williams is the other major protagonist in the contemporary debate about personal identity. Like Parfit, Williams, at one point, draws attention to the fact that the preservation of the numerical identity of a person and the preservation of qualities, which characterize a particular personality, do not always go hand in hand.

He uses a thought-experiment to illustrate this (Williams 1973b, pp. 51-52). In this thought-experiment, Williams asks us to imagine that someone will torture us. He sketches different set-ups for this torture. We should think about whether we would personally fear this torture in all cases. In a first scenario, the torturer tells us in advance that we will lose the memory of the prediction of the torture before the actual torture. Later, he adds that we will in fact lose all our memories before the torture. In a third script, he says that after losing all our memories and before the torture, we will get the impression that we had a different past than we actually had. There is a twist in a last scenario, where the torturer tells us that these new impressions that we will have, will be of events that will already have been experienced by another person. Williams predicts that we will fear the torture in all these cases and that we will in every case think that we are the ones who would be tortured. The fact that we lose all our memories and get other apparent memories instead does not change this. From this, Williams concludes that we behave as if our numerical identity is not constituted by psychological characteristics and continuity, but instead by one particular body.

Yet, like Parfit, Williams soon forgets what is special in this finding. He forgets that we can identify a person as numerically the same by means of her body as token, even though this body does not capture anything about the particular character of a person; i.e. even though it does not capture anything about what we often consider to be relevant about a person. This becomes clear in his analysis of a second thought-experiment.

In this thought-experiment, Williams asks us whether, when we are in love with someone, we would be as happy to continue our lives with

her exact duplicate (Williams 1973c, p. 81). According to Williams, the answer here is 'no.' From this, he justifiably concludes that the body is crucial for the identity of our beloved. We only consider her as our beloved as long as she has the same material history. Williams correctly remarks that we, in this case, follow a person through a body-token and not through a body-type (Ibid). If it would be the body-type of our beloved that made us attached to her, then we should be as happy to continue our lives with her duplicate: this duplicate has exactly the same type of body as our beloved has.

However, from this second thought-experiment, Williams wrongly concludes that it is not "a deep metaphysical error" to say that loving a person and loving a body is basically the same (Ibid). From the fact that we identify a person through her body, Williams derives that this body must also be what we value most in her. This is not necessarily the case. We may identify someone as numerically the same, yet no longer like her because she changed qualitatively. In this case, we would value other characteristics higher than the body, which constitutes her numerical identity. And even when part of our love for someone stems from the particular history that we shared with her – a particular history that no duplicate, but only a being materially continuous with her can have –, her material continuity as such need not be what we love in her. We may love her because she was mentally there for us in the past. In this case, we could judge that she was only there mentally, because she was there physically. So, we could recognize that the real foundation of her numerical identity is physical, all the while, what we value in this person, are her mental qualities, for which her physical composition and history allowed.

In a likely and more favorable reading, Williams's way of formulating things is just unfortunate. By saying that it is not a deep *metaphysical* error to say that loving a person and loving a body is basically the same, he may just mean that, in order for a particular person (for instance, whom we have loved for three years and counting), to still be there, she must be materially continuous with that particular person, whom we fell in love with three years ago. He may mean that in order for someone to still be the same, she must have the same body. If asked about it, he could regret expressing this by saying that we in this case just love a person's body, and reformulate it by saying that, what we here love, is the same person, who can only be the same, if she also has the same body. This would be in line with what he tried to make clear with the other thought-experiment: we can consider ourselves and others as the same person, as long as we have the same body. Unfortunately, he used another, more ambiguous, formulation that expresses the occurring confusion that, what constitutes the numerical identity of persons, must be something that we value in them. Throughout this chapter, I have tried to show that this thought, in its turn, can lead to the mistaken idea that the numerical and qualitative identity of persons must necessarily be constituted by the same elements.

II. WHAT I WILL DO

In what follows, I will continue to distinguish the question of what constitutes the numerical identity of persons from that of what constitutes their qualitative identity. I will examine whether there are

elements constitutive of our numerical identity of which the relevance for our qualitative identity is not immediately evident.

Before embarking upon this exploration, it is important to say something about the kind of exploration to which it belongs. It is not a phenomenology in the Husserlian or Kantian sense. I do not in the first place describe an experience of personal identity as it is always or cannot but be experienced. Neither will it be a primordially conceptual analysis. Above all, it will be an anthropology. I wish to analyze and describe how we factually think about diachronic personal identity.

CHAPTER 2:
THE MENTAL FAILS TO CONSTITUTE DIACHRONIC PERSONAL IDENTITY

My reason for separating the questions of what constitutes the numerical and qualitative identity of persons is twofold. I mentioned the first in the previous chapter: numerical and qualitative identity *are* distinct. We know that objects can be qualitatively identical, while not being numerically identical, as well as that one object can remain numerically identical while changing qualitatively. This is the case for persons as well: identical twins look alike, but are not numerically the same; and a conservative grandfather can be numerically identical to, though not qualitatively identical with, a young revolutionary man of 50 years earlier. Given the distinction between the numerical and qualitative identity of persons, we cannot presume that their constitutive elements are the same. Secondly, it has proven to be difficult to demonstrate that elements constitutive of the qualitative identity of persons also constitute their numerical identity. This makes it interesting to investigate whether the numerical identity of persons cannot have other constituents.

In this chapter, I will describe some of the difficulties that philosophers, such as Derek Parfit and David Hume, encounter in trying to justify that relations that have a crucial role in upholding a person's qualitative identity, also constitute his numerical identity. I will subsequently formulate an alternative proposal as to what could constitute a person's numerical identity.

I. NON-BRANCHED PSYCHOLOGICAL CONTINUITY

Relations, such as those of memory and resemblance, between different instances and expressions of a person's psyche, have a major role in the constitution of her qualitative identity. The less a person remembers, behaves in concordance with the same predictable character, and cherishes the same wishes and projects, the harder it is for us to recognize her as the same character or, in other words, as qualitatively the same person as before. This is, for instance, illustrated by our perception of an elderly person who does not only forget more and more, but who is also angry for reasons and in ways we never witnessed before. About him we may say that it is hard to relate to him as we used to, since not much seems left of the character we once loved. Yet, while memory and the resemblance of psychological character can make or break the qualitative identity of a person, we cannot extend this and say, as Parfit did, that psychological connections constitute the numerical identity of a person – not even if one specifies, like Parfit, that this requires a sufficient degree of continuously overlapping chains of psychological connections, as well as that this continuous stream does not branch. In the remainder of this section, I will explain why this is the case.

1. There is insufficient evidence for us to believe in the existence of psychological *continuity*

To defensibly claim that psychological continuity constitutes our idea of the numerical identity of a person, we should perceive either this psychological continuity, or elements that make us believe that this psychological continuity exists. Yet, as David Hume already observed, there is neither a real connection between all the impressions, thoughts and behaviour that we ascribe to one self, nor is it clear which other relation or apparent relation between these psychological elements could make us *think* that there is such real connection. Let's retrace Hume's reasoning here.

In his *Treatise of Human Nature*, Hume wonders why our imagination forms the idea that we have one diachronic identity.[10]

[10] This is, in fact, a controversial claim. There are two main interpretations regarding the aim of Hume's account of personal identity, as well as his subsequent problem with this account. Don Garrett (1981) and Susan Mendus (1980) hold that Hume tries to determine what selves are, and then finds out that there is a problem with his ontological description of these selves. Galen Strawson (2011) is convinced that Hume is more interested in the working of our imagination and cannot find out what gives someone the *idea* that he is both a synchronically unified and diachronically existing self. Although both Hume's section on personal identity and its appendix begin with some ontological claims about the self and the mind, his text validates the latter interpretation. Hume examines how, given certain ontological facts, the imagination still forms the idea that we are a self with a synchronic unity and a diachronic existence. I will underline his quotes to illustrate this. The italics are Hume's.

The ontological facts are that selves "<u>are</u> nothing but a bundle or collection of different perceptions" (Hume 1978, I.iv.6, p. 253), and that the mind "<u>is</u> a kind of theatre where several perceptions successively make their appearance" (Ibid, I.iv.6, p. 254), but that "[t]here <u>is</u> properly no *simplicity* in it at one time, nor *identity* in different; whatever natural propension we may have to imagine that simplicity and identity" (Ibid, I.iv.6, p. 254).

What follows is this question: "What then gives us so great <u>a propension to ascribe</u> an identity to these successive perceptions, and to <u>suppose</u> ourselves possest

Two observations lead up to this question. First, Hume notices that some philosophers have the idea that each of us is a self who persists through time (Hume 1978, I.iv.6, p. 251). Secondly, he assumes that all ideas are derived from impressions. He looks for the impression of this persisting self, but cannot find it. For Hume, the mind "is a kind of theatre where several perceptions successively make their appearance" (Ibid, p. 254), but in which he can find "no *simplicity* … at one time, nor *identity* in different; whatever natural propension we may have to imagine that simplicity and identity" (Ibid). We do, indeed, have many unrelated thoughts and they change constantly. From these observations, Hume's question follows: if there is no such impression, how then does the idea of being a persisting self originate? Hume continued to believe that we must derive this idea by seeing something in our psychological content. This is shown by his assumption that our idea of being a diachronically existing entity must stem from our idea that there is identity in our perceptions: "What then gives us so great a propension to ascribe an identity to these *successive perceptions*, and to suppose ourselves possest of an invariable and uninterrupted existence

---

of an invariable and uninterrupted existence thro' the whole course of our lives?" (Ibid, I.iv.6, p. 253)

In his appendix, Hume, again, first addresses an ontological question: "Is self the same with substance?" (Ibid, Appendix, p. 635) He says that this question does not really make sense, because, with regard to the mind, we have, in any case, no evidence of self or substance distinct from the particular perceptions. Hume has no problem with this answer of his: "So far I seem to be attended with sufficient evidence." (Ibid). He only worries about how the imagination then forms the idea of a self: "But having thus loosen'd all our particular perceptions, when I proceed to explain the principle of connexion, which binds them together, and makes us attribute to them a real simplicity and identity; I am sensible, that my account is very defective". (Ibid) A bit later, he says "the thought alone finds personal identity" (Ibid) Again, he explains what makes him despair: "But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness" (Ibid, pp. 635-636).

thro' the whole course of our lives" (Ibid, p. 253, my italics)? Hume initially answers his own question by saying that there are different relations, which are not themselves relations of identity, but the perception of which can make us feign that there is an identical object – in this case, a self. He identifies these relations as those of resemblance, causality and contiguity (Ibid, pp. 254-260), but specifies that it is evident that contiguity does not have a part in the constitution of our idea of *personal* identity. According to Hume, the observation of resemblance and the assumption of causality suffice for this particular constitution (Ibid, p. 260).

How do they give us the idea that we are one identical subject with an uninterrupted existence? According to Hume, resemblance and causality are such relations that they spontaneously make our mind associate distinct perceptions with one another (Ibid, p. 254). When we have a perception that resembles an earlier one, this makes us automatically think of this earlier perception. When we perceive something that might be a consequence of something earlier perceived, this earlier event automatically comes to our mind again. Hume states that, because these spontaneous associations make us think back so fast from a later, to a distinct earlier perception, we often ignore the small differences between these perceptions and take them to be similar (Ibid). As images resemble the object they are images of (Ibid, p. 260), we subsequently take them to be images of the same object. When we realize that there was, in fact, a difference between the perceptions we confounded with one another, we try to justify our thought that these perceptions were of the same identical object by feigning that there is either a specific, or still unknown mysterious, connection between what we have perceived that must make these perceptions, perceptions

of one object (Ibid, pp. 254-255). Our idea of identity is, in these cases, a fiction, because it does not correspond to an impression of a real identical object, but is something we feign to justify our thought that qualitatively different perceptions are perceptions of one identical object. As far as the fiction of personal identity is concerned, Hume attributes a major role to memory in its production and discovery. He says that memory produces this fiction, because it produces the resemblances between perceptions, which makes us think that they are completely similar and must thus be of the same object:

> "suppose we cou'd see clearly into the breast of another, and observe that succession of perceptions, which constitutes his mind or thinking principle, and suppose that he always preserves the memory of a considerable part of past perceptions; ' tis evident that nothing cou'd more contribute to the bestowing a relation on this succession amidst all its variations. For what is the memory but a faculty, by which we raise up the images of past perceptions? And as an image necessarily resembles its object, must not the frequent placing of these resembling perceptions in the chain of thought, convey the imagination more easily from one link to another, and make the whole seem like the continuance of one object" (Ibid, pp. 260-261)?

Hume then assumes that we confabulate personal identity when we experience something of which we later also have memories. The reason is that these experiences and memories seems so similar that this makes us imagine that they must be of the same object, i.e. of the same perception. According to Hume, our memory further discovers personal identity, because it allows us to recall two distinct ideas and so to imagine that there is a causal link between them and that they must thus belong to one mechanism (Ibid, p. 261). This explains why we can judge that a person is still numerically the same, even when his character changed. We could recall an event that made *him* change

character and so still think of him as the same person. Once memory lets us discover that a person can be the same, even though we have a different perception of him, as long as the difference in his appearance can be explained by a cause that made him change, this can also make us imagine that we were a person of whom we no longer have any memories (Ibid, p. 262). For this, it suffices that we imagine that there are causal links between how we are now and how we were then. How we think and behave now, could be explained by how we thought then, and by what we experienced in between.

Yet, in the appendix of his treatise, Hume rightfully admits that something is missing in his analysis of the constitution of our idea of personal identity, if it is to be an analysis of what constitutes our numerical diachronic identity: it does not explain what motivates the imagination to connect certain impressions in such a way that this makes us assume that there is one persisting self (Ibid, Appendix, pp. 635-636). The fact that we perceive resemblances or assume causality between different impressions does not really explain this, because we also assume that all these impressions could occur without being related to one another (Ibid, p. 634 and p. 636). It is not because I have an impression at time 2 that resembles an impression at time 1, or could have been caused by an impression at time 1, that my impression at time 2 depends on my impression at time 1. I could have the latter impression without having the former. Something more is needed to make us think that these two events are indeed connected to one another. But in what this something more could consist, remains unclear.

There is nothing in the propositional content of thoughts that refers uniquely to one person. In principle, different persons can have

the same thought, propositionally speaking. Nor do the different thoughts one has necessarily induce each other: one specific thought can lead to many different thoughts, or just not induce another thought at all.

Following one self through his character alone is impossible for similar reasons. First, different persons can have the same character. Secondly, it is not possible to entirely predict how one character will develop and thus to derive from one or more character traits which past or future person someone is numerically identical to. We are not cartoon figures who stick to a specific role. There are different possible ways for us to develop ourselves and we can suddenly do something unexpected.

Nor are there other promising psychological relations that could themselves, either connect all the thoughts and conscious experiences we ascribe to ourselves, or give us the impression that they are connected in this way. The often suggested relation of memory is no such relation. We ascribe more experiences to ourselves than we can ever reach through memory. We, for instance, say that we were once a baby, dreamt a dream to which we no longer have access, or fainted. All the while we had experiences as a baby, in dreams, and moments before we fainted, which we *never* remembered – not even right after we had them.

The feeling of mineness of experiences has also been suggested as an aspect of our experiences that can make us ascribe a particular set of experiences to one self. Every experience which someone feels to be his, would then be his. The Danish philosopher Erich Klawonn (1987) held that this feeling is what individuates us. He argues that if we imagine that we step into a machine and two completely similar

persons step out of the machine, of the similar persons, we would claim to be the one with whom we share the feeling of mineness. In other words, according to Klawonn, I would be the one who experiences his experiences as his, just as I now experience my experiences as mine.

Klawonn's reasoning is problematic. It leaves a question unanswered. How could one compare the feeling of mineness one once had with the feeling of mineness one currently has and, from this, derive whom of the two duplicates one is? Such a conclusion cannot be based on this feeling alone, since the feeling of mineness of experiences is the same for both duplicates. All subjects, whom we call mentally healthy, experience their experiences as their own. Both duplicates will thus experience their experiences as their own. But that would be all that there is to it. This feeling does not have any other particular colour that is identical to, and individuates, the feeling of mineness I had before stepping into the machine. My feeling that my experiences are mine does not feel different than how the same kind of feeling feels to others. I can, therefore, not ascribe an experience to one self or even to myself, just because it felt like being the subject's, in a similar way to how my feelings now feel like they are the subject's, i.e. mine. It is no solution to say 'as long as I still experience a feeling as mine, it is mine' – and to so avoid having to compare two distinct feelings of mineness, i.e. a current and an earlier one. For, as I described above, we ascribe experiences to ourselves to which we no longer have access, not even through memory. If I cannot access them, I can neither access their taste of being mine.

Given that it remains unclear in what psychological continuity could consist, we cannot defensibly maintain that the observation of

real or apparent psychological continuity constitutes our idea of being one diachronic self.

2. Our impressions, thoughts and behaviour do not all belong to one continuous psychological stream

Apart from it being difficult to see in what the psychological continuity, which would constitute personal identity, could consist, psychological *connectedness* and *continuity* are not what we should be looking for, when we examine where our idea of personal identity comes from. The claim that a psychological continuity made of an at-all-times sufficient degree of interconnected psychological connections is necessary, and, with the criterion of non-branchedness added, sufficient for the constitution of numerical personal identity does not account for how we actually ascribe personal identity. In real life, we are prepared to consider a being as one numerically identical person, even when he has thoughts that are in no way connected to one another and that can thus not belong to one psychologically continuous chain. We constantly have such thoughts. One example can be that of a person in mourning who suddenly sees something funny that makes her laugh. Her observation of whatever was comedic is then in no way connected to the thought that makes her mourn. Still, we attribute both the sadness and the comic experience to the same person. We do so without feeling that we should, in principal, check whether there are other psychological links between the person who laughs and the one that mourned just a moment ago.

3. We attribute numerical identity to persons with branched streams of consciousness

Lastly, *non-branching* psychological continuity is not what we should be looking for either, when we examine what makes us re-identify someone as numerical identical. Parfit reasons that the numerical *identity* of a person is lost as soon as his earlier stream of consciousness branches (Parfit 1984, pp. 256-257). If branch B and C branched off A, they would start to consist of distinct experiences and thus not be identical to one another. Consequently, we can neither say that they are identical to the original stream A, with which they are psychological continuous. For, if B is identical to A, and C is identical to A, then B and C would have to be identical to one another, but this is not the case. Locke, too, claimed that the consciousness, with which we remember something now must be the *same* as the consciousness, with which we first experienced the event (Locke 1996, II.xxvii.9, p. 138). Yet, two problems arise here.

First, it is unclear how psychological elements themselves could guarantee that a certain psychological continuity will not branch. Parfit admitted that they cannot. According to Parfit (1984, pp. 199-200), we could theoretically copy one psychological history, transfer this to two different bodies and then let them each further develop in their own way, giving rise to two different entities. Parfit also refers to the real case of split brain patients. These are patients, of whom the connection between their hemispheres was cut and whose consciousness produced by the one hemisphere subsequently has no access to the consciousness produced by the other. One could say that,

what Parfit calls 'psychological continuity', branched here: there was one psychological 'continuity' before the cut and two branches of this 'continuity' after. If psychological continuity cannot guarantee non-branching and non-branching is a criterion for personal identity, then there must be another element than psychological continuity itself that helps to constitute personal identity.

This is, however, not the conclusion that Parfit draws. He maintains that personal numerical identity only stays preserved when there is an unbranched psychological continuity, which is made up of a sufficient degree of psychological connections and does not add another element (outside of the psychological realm), which would have to prevent this continuity from branching. Instead, he argues that the continuation of personal *identity* is not what really matters to us when we want someone to survive. What we value in a person's survival are the continuation of his projects and character, which his psychological continuity makes predictable (Parfit 1984, pp. 261-263, 271). Given that this psychological continuity never guarantees non-branching and thus numerical identity, it would be rational for us just to care about the continuation of a psychological stream, regardless of a numerical identity, which it in, any case, does not guarantee. I disagree with Parfit's conclusions in this matter. First, identity does matter to us. We sometimes have the wish to finish projects *ourselves*, as well as the wish that others live long enough to personally witness how one of their dreams comes true. Secondly, Parfit's assumption that non-branched psychological continuity is what we look for, when we examine what could constitute personal identity, is mistaken. This is also the second problem with advancing non-branched psychological continuity as a criterion for the numerical identity of persons. In

reality, we do attribute branched streams of consciousness to *one* self. This emerges, for instance, from the experiments that Gazzaniga performed on split brain patients.

In the beginning of the fifties of the previous century, R.W. Sperring and Ronald E. Myers discovered, how when the cerebrum is surgically divided by cutting the corpus callosum connection between the two hemispheres, the brain seems to contain two separate spheres of consciousness. When one performs tests on a subject with such a brain, one can see how the left hemisphere processes information, to which the right hemisphere has no access, and vice versa. Michael Gazzaniga (1967) gives some clear examples of this in 'The Split Brain of Man.' Together with Sperring, Gazzaniga examined this phenomenon in human subjects, whose connections between the two hemispheres were cut in a (successful) attempt to reduce the impact of an epileptic seizure. They first presented ten of these patients with a ray of light which passed through their left visual field (where the information is processed by the right hemisphere), and through their right visual field (where the information is processed in their left hemisphere) (Gazzaniga 1967, pp. 24-25). The patients reported that they saw light flashes when these passed through their right visual field, yet denied that they saw them when they passed through their left visual field. At first sight, it might seem as if these patients are partially blind, but this proved not to be the case. In fact, the patients also saw the flashes in their left visual field. When asked to point them out, they could. They were just unable to verbally report them, because the information about these light flashes was processed by their right hemisphere, while their ability to speak is mainly governed by their left hemisphere. The following situation then presents itself: when

light flashes appear in the patient's left visual field, he may verbally deny that he sees them, while pointing them out. There then appear to be two separate streams of consciousness: one, in which there is awareness of the flashlights and the will to point them out, and another, in which there is no awareness of the flashlights and the will to deny that any flashlights are seen.

The relevant point of this phenomenon for what I previously said is what follows. When these patients are confronted with their radically opposing answers, they try to explain them in such a way that they can still appropriate them both. One case, described by Gazzaniga, (1967, p. 29) is that of a woman whose right hemisphere made her chuckle after it made her see that the experimenters suddenly let a nude appear in her left visual field. She would never *say* that she saw this nude, because her left hemisphere, which is the one that allows her to say anything, did not make her see this nude. When the experimenters asked this woman why she chuckled, she therefore hesitated for one moment, but then said 'because of this funny machine', indicating a machine that was shown in her right visual field. The hemisphere which makes her say that this machine is funny, in fact, never made her experience this machine as funny. Still, the lady assumes that she thought that this machine was funny, seemingly because she assumes that the body, which chuckled, was hers and that she must thus have been chuckling about something. Split brain-patients then seem to appropriate experiences because they were processed by one body that they consider to be theirs, not because they have an inner mental access to them, and regardless of whether this implies that they are the subject of multiple branched streams of consciousness.

This is confirmed by another case described by Gazzaniga (Ibid, p. 27), in which the experimenter asked a patient, who processed a red light flash with his right hemisphere, what the light flash was like. The patient had to use his left hemisphere to answer verbally, but this hemisphere had not processed the light flash. So the patient guessed and said that the light flash was green. His right hemisphere made him hear what he just said and realize that this was wrong. As a consequence the patient frowned. Now, the patient's left hemisphere made him observe this frown. As a result, the patient corrected himself and said 'no, I saw a red light.' Here again, the patient did not ascribe an experience to himself because he had an inner, private mental access to one unified stream of consciousness, but rather because he observed a meaningful reaction of a body he considers to be his.

If psychological connections, which are important for the qualitative identity of persons, cannot constitute numerical personal identity by forming a continuous and unbranched chain, which would be necessary to constitute the diachronic and unique existence of this one person, one or more other elements must help constitute our idea of such an identity. In the remaining part of this thesis, I will argue that our body, considered as mere material token, has an important role in this constitution: we only call those persons numerically identical who live through one continuous body. In the remaining part of this chapter, I give some first factual evidence in support of this thesis; enumerate some elements that make the body – compared to psychological continuity – into a more evident constituent of the numerical identity of persons; and indicate, in which kind of self-consciousness we could indeed identify a self through its materially

continuous body. The coming chapters will further extricate the here just briefly forwarded thesis.

II. A ROLE IN THE CONSTITUTION OF OUR NUMERICAL IDENTITY FOR OUR EXISTENCE AS MATERIAL ENTITIES

1. Factual evidence

There is factual evidence that supports the claim that we judge that persons who live through one continuous body are numerically identical. For instance, if a never-convicted World War II Nazi officer is now discovered, but he happens to have amnesia and he happens to be a very kind person now, then we may choose to no longer legally convict him for the crimes he committed as a young officer. Still, newspapers and biographers of the person retrieved will write that *he* once was a Nazi officer and now suffers from amnesia. All that links him to this officer is the body he has now: it developed out of the body of the Nazi officer, i.e. it is materially continuous with it. Apparently, in biographies (and autobiographies alike), we include, as actions in the subject's life, everything done with the consciousness produced by a body continuous with his current body; even, and often especially, that which is foreign to his general way of behaving. This does not change when the current subject can no longer identify himself with these actions, or no longer remembers them. Family members of amnesia patients can treat these according to a similar logic. The personality

of some amnesia patients changes in such a way that some relatives decide to no longer communicate with them. They stop rendering visits to a person who they no longer recognize. Still, these same relatives would notify their acquaintances of the fact that their mother, or a particular person with the name 'X' died, when this amnesia patient died. They then assume that the person they once knew and loved was somehow still there until her last breath – even if she was, at that point, just materially continuous with the person she was before and no longer psychologically similar or connected to this earlier person. Other examples indicate that the same can be the case when we attribute identity to ourselves. When my mother shows me a beautiful painting that, according to her, I made as a child, I can feel proud of it, as well as regret that I have lost a talent I once had. I then acknowledge that I once was this painter, even if there is no psychological connection between my current self and the earlier painter. I can acknowledge that I am a person, on the basis of the testimony of others who witnessed, from the outside, how I was physically present at a particular past event. Just like this, I can acknowledge, on the basis of video-images, that I have broken a vase when I stumbled drunk into the house, even if I had a blackout soon after it and do not currently remember this event.

2. Characteristics of the body that render it suitable to constitute the numerical identity of persons

The body also has certain characteristics that render it more suitable than the psyche to function as a criterion for the numerical identity of a person. A first advantage of the body, in this regard, is that it is clearer in what its continuity exists. At one particular time, it is always at one particular place and, if you trace it, there are no gaps between these places. Even if you are not always aware of this body or its place, you know that it is somewhere. A machine could track it constantly and tell you where it was at any given time. At any one moment in time, all its actual parts are physically connected to one another, meaning that they are either immediately attached to one another, or through other cells, which are themselves part of the body. Throughout time, some cells of the body die, while others are generated. Yet, if we would film the body uninterrupted with microscopic cameras, we would see that every cell it has, developed from, or is kept alive by, other cells which are part of the body, i.e. by cells that are themselves organically and physically connected to the other cells of the body. I mention 'or are kept alive,' to allow for the possibility that transplanted tissue or organs can also become part of one continuous body. A former, or later, body part can, at one point, be organically and physically separated from a physical body, of which it is a part at other times, but it is only part of one continuous body for as long as it is still physically attached to it and organically kept alive by other parts of it. Of course, our observation of physical continuity may be fallacious. We can think that we observe physical continuity

because we see a particular body now, at a place very close to the place where we saw a similar body one moment ago, all the while the second body did not really develop from the first, but just seems to do so because it came into existence soon after, at a place nearby. In this case, the appearance of continuity would still explain why we think to observe continuity and so assume identity. We visually perceive the contiguity of this body: we do not actually see its continuity (in principle we could be confronted by different quick body flashes and just think that we perceive continuity), but we can follow an at-least-seemingly continuous body (when we keep our eyes open, we do not see it disappear and come back again and we see how it only displaces itself to an extent, which we take such a body to be capable of performing in a particular time). Thus, our impressions can make us think that we have a same continuous body in front of us.

A second reason to find it more credible that our bodily continuity is crucial in our attribution of numerical identity to persons than is our psychological continuity, is that this accounts better for the fact that we ascribe unconnected thoughts to the same self. If we attribute numerical identity to a person as long as he consciously lives through the same continuous body – it being unimportant what he is conscious of –, then he could still be the same self even if some of his thoughts seem to emerge from nowhere and not be connected to any other thoughts he has.

Thirdly, if we say that a person stays numerically the same, as long as he lives through one continuous body, this concords with the phenomenon that we can ascribe two different streams of consciousness to one self. The fact that bodies, too, can split (cfr. identical twins), or branch (cfr. Siamese twins), does not endanger

their suitability for being a crucial criterion for the numerical identity of persons. The non-branch requirement for the preservation of the numerical identity of persons would still be respected. When a human body that generates the life of conscious being branches, or a fortiori splits, we see the resulting entities as two different persons. One could discuss what counts as branching in this case, but there seems, at least, to be a clear consensus that analyses the obvious example regarding this topic, namely, that of Siamese twins. We consider these twins to be two persons, when they are blessed with two heads and respective thought-producing brains. However, when the newborn does not have two heads, but only more than the usual limbs, we will judge that there is just one baby and that his twin brother never was or died. Still, it is the branching of the body, and not of the psyche, which here makes us consider the conscious being in front of us as two persons, rather than one. If a certain and critical mass of the body branches into two physical branches, which each still generate consciousness, we judge that we have two persons in front of us. Whereas the same is not the case when a one headed human being verbally denies, but, through writing, confirms that he saw something. We then have two branches of consciousness, but still assume that there is only one person. So, if a consciousness-producing body branches in a certain way, we assume that there are two persons, whereas if consciousness branches, we can still assume that there is one person. If we then logically hold that what constitutes numerical identity cannot branch or split without destroying this *numerical* identity – per definition only attributed to one entity –, only the body, and not consciousness, can function as this constitutive element.

If the continuity of our bodies is indeed key in the constitution of our idea of what the numerical identity of a person consists in, then Hume could incidentally appeal to his own discourse to resolve the problem with his initial account of the origin of our idea of personal identity.

As we saw, Hume wondered why we assume that particular resembling impressions belong to one person and why we think that there should be a causal link between particular perceptions. This is unclear, because of the combination of the facts that (1) we do not perceive a real connection between these impressions and (2) we could have every one impression without a particular other. It will be explained, says Hume, if our perceptions would "either inhere in something simple and individual" (Hume 1978, Appendix, p. 636), or if the mind did "perceive some real connexion among them" (Ibid).

These references to something real may surprise one, since Hume's question is here not whether there really is a self, but instead how we get the *idea* that we are diachronic selves. It is not immediately evident why Hume expects that something real may have a part in the genesis of this idea. After all, he was the one who made us skeptical about whether our ideas correspond to something real. Hume (Ibid, I.i.6, pp. 1-7) argued that we can, at most, find out whether a particular idea corresponds to a specific impression and shows, with the case of identity, that even this does then not always prove to be so. According to Hume, some ideas, like those of identical objects and subjects, are clearly fictions to which no particular impression corresponds, but which originate after an association of different perceptions (Ibid, pp. 254-255). There does then not seem to be a role for something real in the origination of these ideas.

Yet, Hume's reference to the real could be explained in the following way. We may understand conceptually what it would mean for our perceptions to reside in something real or to have a real connection between them. This understanding could then make us associate specific perceptions with one another. This interpretation does not deny Hume's skepticism. For it is still possible that when we get the impression that this real entity or connection is there and we associate those perceptions which reside in this apparent entity, this entity does not in fact exist. Our impressions might mislead us about the reality of this entity, all the while our thought that it might be real, may make us associate those impressions with one another which seem to reside in this entity.

With this insight, Hume could turn to a phenomenon which he observed, but subsequently deemed of no importance for the constitution of our idea of diachronic selves,[11] i.e. that an observation of contiguity can make us think that something has a diachronic identity. Contiguity, for Hume, is the nearness in time and place of objects that can make us think that they are identical, and so, continuous. Hume did not see how this could be of relevance for the constitution of a self that he most likely considered to be a mental entity. Yet, our perception of almost the same bodies, close to each other over time, could make us think that there is one real particular body. This conception of one apparently real entity could then make us associate those impressions with one another which seem to be produced in or through this body. The latter association of certain

[11] "And here 'tis evident we must confine ourselves to resemblance and causation, and must drop contiguity, which has little or no influence in the present case" (Hume 1984, I.iv.6, p. 260).

impression could in its turn make us think that there must be one person who has all these impressions. If, at first sight, unimportant elements for the qualitative identity of a person can constitute their numerical identity and the contiguity of the body appears to have a role in the constitution of our idea of being a diachronic self in the way just described, then Hume could find the real connection he was looking for in the apparent continuity of the body. This apparent continuity or 'contiguity' is real because it is there, regardless of our individual subjective experience of it. It can be registered by others or machines.

3. A conception of a self which corresponds with tracing it through its body

If the non-branched material continuity of a body that generates consciousness is indeed a crucial criterion for the numerical identity of a person, then what a person is like may not have a role in our acknowledgment that he is a particular self. This would be in line with Hume's findings that we do not derive the idea of being one self from the observation of resemblance or causality between thoughts. Our idea that we are diachronically existing beings could precede these observations.

In what follows, I will argue that one of our ideas of being a self is indeed rather formal. I can work with the idea that I am one diachronic entity without first having to think about which particular one character I have. One example of this is the reference I make to myself

in an ordinary conversation. When I talk to a person and say 'I think this,' I can correctly refer to myself without first having to think about which character this 'I' has.

Our formal idea of being a diachronic self comes about through a self-consciousness that is not regularly mentioned in the discussion about the constitution of personal identity. Contrary to the consciousness mostly referred to, it is a consciousness of one self that does not come about through a reflection on the contents of one's own consciousness. I argue that our ascription of numerical identity to persons is an ascription of identity to them as formal selves and occurs as long as they are (dispositionally) conscious beings living through one materially continuous body. Here are some examples to support this claim. When someone asks me how old I am, I usually answer without hesitation. If I have to think a little, I think about my previous birthday and then remember that this was my 25th, or I remember the fact that this is the last year that I will be able to purchase youth tickets for the train and that this is so because I am 25. Here, I reason about myself, while no particular image of this self crosses my mind. I do not think about which character I have or which thoughts I have had. I just take it that I have been here since the day I was born and that I am as old as my body is. Another example is the following. We save money on our name, which gives a bank clerk the right to later return it to us. This clerk does not identify us by our face or smell, as an animal might recognize another animal. He identifies us by our first name (which is a sign for us, but not a pictogram), the name of our parents, and our place and date of birth. All of this does not only require that we are selves, but also that we have the idea that we are a diachronically existing self and deal with other such selves. Yet, here

again, we just work with the idea of being a particular self, without therefore first having to think or be convinced that this self has a specific character. If we would reflect on our behaviour, then we would see that we, in this case too, assume that we are the same as long as we are bodily continuous: we expect assurance from the bank clerk that we get our money back. He should check whether the person who requests it is physically identical to us, not whether he or she thinks like we used to do. It is true that there is a rule that gives the clerk the right to allow that our heirs manage our account when we are no longer lucid. Yet, even this confirms that we follow ourselves through our bodies: heirs are the persons who were appointed as heirs by the lucid self which was connected to the body to which now only a self with amnesia is connected. They are not appointed by a self who now thinks how I used to think. These examples illustrate how we can think of a self in a rather empty way, as well as that we identify this self through a particular body.

4. What is next?

I have now analyzed why our thoughts, feelings and psychological behaviour cannot constitute the numerical identity which we factually ascribe to a person. I have also given initial plausibility to the idea that our acknowledgment that we live through one continuous body may have a part in the constitution of our idea of what makes a person numerically identical. I gave some examples of how we factually attribute numerical identity to persons, mentioning characteristics of the body that make it suitable as constitutive element for this identity.

I further illustrated that we – amongst other conceptions of the self – work with a rather formal concept of a self that could be tracked by tracing the body it lives through.

Until now, nothing radical should be concluded from the above. So far, we can only conclude that, if persons were essentially mental, it would be hard to identify and re-identify them. This would be easier, if having one materially continuous body is part of what characterizes them. So far, I only considered the pragmatic advantage of the body in identifying and re-identifying persons. In the fourth chapter, I will argue for the further claim that, given that persons are particular re-identifiable entities, they cannot merely be mental entities, but must have a material body.

Before arguing for this further claim, I will use the third chapter to elaborate on the specifics of my examination of the role that the body may have in the constitution of our numerical identity. Contrary to some other theories that attribute a role to the body in the constitution of persons, I will not argue that our body has a role in this constitution because we are embodied subjects, who feel and act intelligible through their respective bodies. I support this idea, but will not elaborate on it in this dissertation. Instead, I will analyze what the function is of the body in the constitution of our numerical identity, when we consider it as mere material entity or token. A second difference between my approach of the role of the body for personal identity and those of many others is that I do not wish to demonstrate what the body tells us about who we are. I will not argue that the body is meaningful and can tell us who we are as kind, or as particular person. I will argue that the body has a crucial role in how we ascribe numerical identity of persons, but this will not imply that the body gives a specific sense to this

persons. The body will prove to allow people to develop a specific and meaningful character, but the meaning which persons have for us, is not had by this body.

In the fourth, fifth and sixth chapter, I will examine what could explain the role of the material body in the constitution of our numerical identity, if not an apparent importance of this body for someone's personality. I will analyze which are the properties of material bodies and of our self-consciousness that incite us to identify a self through a body.

In the seventh and last chapter, I elaborate on which image of the self arises, when we really follow a particular self through a particular body.

CHAPTER 3:
THE ROLE OF THE BODY HERE CONSIDERED

In the previous chapters, I have distinguished the question of what constitutes the numerical identity of a person from that of what constitutes his qualitative identity. I suggested that the body qua material token may have an important role in the former constitution. This is a specific stance that differs from other theses forwarded about the role of the body in the constitution of personal identity. While this role has been extensively discussed by philosophers working in distinct fields, this is less the case for its role qua material token. Below, I will distinguish my discussion of the role of the body in the constitution of (our idea of) personal identity from those of animalists and protagonists of the constitution theory. The goal is not to enumerate all theories concerning the role of the body in the constitution of personal identity – there are many more –, but instead, to bring out some special characteristics of my discussion of this role of the body, as well as of this body as material token – the aspect of the body that I consider. I first sketch what the animalistic theory, as developed by Eric Olson and Paul Snowdon, as well as the constitution view of Lynne Rudder Baker entail. Subsequently, I enumerate three of the major differences between the functions of the body that are highlighted in respectively these accounts and mine.

I. ANIMALISM AND THE CONSTITUTION VIEW

1. Animalism

Snowdon shortly explains the animalistic position as follows:

> "There is no real controversy…over the claim that certain continuities to do with an animal's body are sufficient for the persistence of the animal. If the body of an animal remains intact and sustains the processes we call 'life', the animal in question has survived. Animalism seems to imply that such conditions are sufficient for *our* survival" (Snowdon 1995, p. 71).

Olson more clearly puts that, according to animalists,

> "no sort of psychological continuity, with or without further physical qualifications, is either necessary or sufficient for us to persist through time" (Olson 1997, p. 4).

Animalists claim that

> "…all *human* people are animals. We are what Locke called 'men'. You are not merely intimately connected in some way with a human animal. The claim is not that your *body* is a human animal, or that you are 'constituted by' an animal. That living primate sitting in your chair right now is you: you are numerically identical with an animal" (Ibid, p. 17).

Since

> "psychological continuity is neither necessary nor sufficient for a human animal to persist through time….the animal that survives the loss of its mental properties is you, if you are an animal, and so you can persist without psychological continuity of any kind" (Ibid).

Animalists admit that

> "[p]erhaps we cannot properly call that vegetating animal a *person*, since it has none of those psychological features that distinguish people from non-people (rationality, the capacity for self-consciousness, or what have you)" (Ibid).

But they say that

> "[i]f so that simply shows that you can continue to exist without being a person, just as you could continue to exist without being a philosopher, or a student, or a fancier of fast cars" (Ibid).

2. The constitution view

Lynn Rudder Baker, protagonist of the constitution theory, agrees with the animalists in that the body has its part in the constitution of a person: "…a human person is constituted by a human body" (Baker 2000, p. 3). According to Baker a human person is "necessarily embodied" (Ibid, p. 6). This means that "one could not exist without having some body or other; but it does not follow that one must have the body that she in fact has" (Ibid, p. 6).

However, "constitution cannot be understood as identity" (Ibid, p. 27) and Baker precisely differs from animalists in her claim that "a human person is not identical to the body that constitutes her" (Ibid, p. 3). According to Baker, persons are not identical to animals because they have some properties that are not animal properties. Persons can have these properties because they have a first-person perspective, which animals do not have: "We human persons are animals in that we

are constituted by animals, but, having first-person perspectives, we are not 'just animals.' We are persons" (Ibid p. 4).

Following Baker, the persistence condition of a person depends on the continuation of this first-person's perspective:

> "If the Constitution View is correct, then what I am most fundamentally is a person, a being with a first-person perspective; and I would cease to exist if that first-person perspective were no longer exemplified. As a human person, according to the Constitution View, I am constituted by a human body; even so, my continued existence depends on the continuation of my first-person perspective" (Ibid, p. 6).

What is this first-person's perspective and which unique person properties does it allow for? In Baker's vocabulary, having a perspective is not the same as having a first-person's perspective. "Many nonhuman animals have perspectives (determined perhaps by the positions of their sensory organs – e.g., eyes that are the sources of their visual fields)…" (Ibid, p. 21). Yet, only persons have a first-person's perspective, which means that they have "a conception of [themselves] as being the source of the perspective" (Ibid, p. 21). To have such conception, "it is not enough to *distinguish* between first person and third person…" (Ibid p. 64). Animals, which are not persons, can make this distinction. A dog, for instance, feels whether it is he, who is angry, or rather the other dog: he might bite in the first case and walk off with his tail between his legs in the second. To have a first-person's perspective "one must also be able to *conceptualize* the distinction, to conceive of oneself as oneself" (Ibid, p. 64).

The fact that a first-person's perspective is unique to persons has as a consequence that

> "…persons are the only beings with the ability to ask the question 'What am I?'" (Ibid, p. 6) "A person [also] has causal powers that a body would not have if it did not constitute a person. For example [those of] using a passport, voting in local elections, and being responsible for a prisoner" (Ibid, p. 109).

An additional unique property to beings with a first-person's perspective is

> "what we might call 'interiority'. By 'interiority', [Baker means] those aspects of ourselves that we can report but that are not directly observable by others. …Any individual who has ever imagined himself or herself in some nonpresent circumstances has an inner life…..To imagine oneself receiving an award or being arrested for speeding is to exercise a first-person perspective" (Ibid p. 161).

II. THREE DIFFERENCES

There are three major differences between these accounts of the role of the body in the constitution of the person and my discussion of this role in this dissertation. (1) In contrast with animalism and the constitution view, I am more interested here in what constitutes our *idea* that we are diachronically existing persons, than in what makes that we *are* these ontological entities. (2) Secondly, contrary to animalists and defenders of the constitution view, I do not highlight in which respects our body has a role in the constitution of our *personhood* (by making us alive or conscious). Instead, I examine what could be the role of our body in the constitution of our (idea of) personal identity if we consider this body as pure material token – an aspect of the body whose importance for the constitution of personhood is not apparent.

(3) Lastly, I do not merely consider how (our idea of) personal identity can, in part, be constituted by our body as an ontological entity, but also by this body as a bearer of meaning.

1. Ontological theories

Animalism and the constitution view are ontological theories.[12] They intend to say something about what things *are*. Olson (1997) defines us as a specific kind of living organism. Baker (2000) argues that we have some person, animal and material properties. In contrast with this, I seek to determine what makes us *assume* that someone is still the same person, rather than what, ontologically speaking, makes a person the same person.

One of the consequences of this distinction between our respective research questions is that I can, just like Olson, hold that we are animals (and that many of our ideas about persons express this), while I can, in contrast with Olson, at the same time indicate that we rather identify a person through his body (considered as material object), which sustains our life functions, than we follow ourselves through these life sustaining functions themselves. This could be so, in spite of

[12] "All of these major competitors – Immaterialism, Animalists, and my own Constitution View – are competing ontological answers to the question 'What am I?' They are ontological answers because each competitor purports to say what most fundamentally I am (and you are) and to give conditions under which I (and you) continue to exist" (Baker 2000, p. 5).

the fact that we would not be the same self, if our metabolism and other life-sustaining functions came to an end.

Another consequence of my diverging research question is the following. Olson and Baker hold that we can find out what the meaning of a person is by looking at his intrinsic characteristics.[13] In contrast to this, I contend that a person's meaning is not solely constituted by what is internal to him. Persons are capable of establishing part of their meaning – i.e. of what they are – through a reference to something external to them. So, even when persons are not identical to their material bodies, these bodies may have a role in their constitution. An example. A conscious person could realize that it is his body that usually makes other people recognize him as the same person. This can prompt him to seek a connection between certain mental experiences, which occurred to whatever conscious being lived through this body: he may, for example, be interested to see whether there is a relation between the thought that a thirty-year

[13] At one point, Baker seems more nuanced than this. She writes: "…once we see that a thing can have relational properties, essentially, we are freed from supposing that if *x* and *y* differ in primary kind, then there must be an actual physical intrinsic difference between *x* and *y*.…What grounds a difference in primary kind is difference in essential properties, whether intrinsic or relational" (Baker 2000, p. 196). However, in fact, she reduces these relational properties to intrinsic properties: "Rivers and aggregates of H_2O have different essential properties and hence are of different kinds….the constituted thing differs from the qualitatively similar constituting thing because of this: There is some intrinsic property… that the constituted thing has essentially but the constituting thing has accidentally, or vice versa. And this difference in essential properties grounds the difference in primary kind" (Ibid, p. 171).

old with this body has now and the thoughts that the teenager with this body had some fifteen years ago. When looking for and expecting such a connection, he is likely to find one. This may change and strengthen his sense of identity. He may never have come to this sense of identity, if he did not have the idea that others followed him through his material body.

The fact that we often recognize someone through his material body may not only have an effect on how we conceive of our own identity. It could also have an effect on how we conceive of someone else's personal identity. Specifically, it could have as effect that when the being who lives through a particular human body suddenly establishes another character than the being who lived through this body before, or when this being, contrary to this previous being, does not seem to have any mental capacities, we will not immediately say that we now have to do with another person. Instead, we may try to interpret this second being's words in the light of what we know about the first being's life: we could wonder whether this second person is saying what the first person said, because the first person had a brain hemorrhage; we could also try to find out whether the first being perhaps already fostered a hidden anger, which the second being now exuberantly expresses. In the case where the second being, contrary to the first being, does not seem to have any mental capacities, many of us may, nevertheless, not easily let go of him. When we are used to identify persons through their bodies, we may think that the being in front of us is in fact the person we knew earlier, even when other characteristics typical of a, or the same, person are lost.

With regard to the ontological analysis of personal identity, my view may, at first sight, seem to be more related to the animalism of Olson and Snowdon, than it is to Baker's constitution view. Indeed, I repeatedly stress the role of the body in the constitution of our (idea of) personal identity. As seen in the previous chapter, I also point out some difficulties with more psychological accounts of personal identity. However, when I emphasize that the body has a role in the constitution of our (idea of) personal identity, I do not wish to decline that a first-person's perspective equally has a part in this constitution. Looking at how we typically talk about persons, and so express how we think about persons, there are at least two cases which support that a first person's perspective, as defined by Baker, is important for the constitution of personal identity.

(1) Consider two human beings shortly before their passing away. One of them is in a vegetative state: he breathes and his heart beats, but he has no idea of being a self. The second being has a first person's perspective. She makes all kinds of outrageous claims about who she is, none of which align with the reality the person associated with this human being was conscious of earlier. After their passing away and in commemorating the person previously connected to the corpse we now greet, we will, in the first case, say that he was, in fact, already gone for a while, whereas we will say, in the second case, that *she* had become unrecognizable at the end of *her* life. The remaining first-person's perspective of the human being in the second case makes us assume that the person who we always associated with this human being was still there, in spite of the fact that her character changed.

(2) We make a clear distinction between a person who suffers from locked-in-syndrome and is fully conscious, albeit he has no means to

express this, vis-à-vis a comatose human being in a vegetative state. When we remove their respective breathing aids, the conscious state of the first person will make us call this murder, whereas the unconscious state of the second being will make people differ over whether she has been murdered or not.

This is why I refrain from saying that persons are just human animals, as Olson and Snowdon claim, and will instead defend that a person is (considered as) the same person as long as his same continuous body produces consciousness.

Yet, my view on personal identity does not exactly coincide with Baker's view either. Baker holds that we could, in principle, survive when there is no spatio-temporal continuity of body:

> "Although I am doubtful that such a procedure is empirically possible, I do not want to take issue with the point that … psychological continuity does not require continuity of body or of matter of any sort. It is thus possible that there is psychological continuity throughout an interval during which there is no spatiotemporal continuity of body" (Baker 2000, p. 126).

If having a continuous first-person's perspective falls under psychological continuity, this means that, according to Baker, we could survive without spatiotemporal continuity of body. I will defend the opposite: a person is (considered as) the same person as long as the same continuous body produces consciousness. According to the account which I will develop, the spatiotemporal continuity of the body is thus as indispensable for the survival of a particular person, as is the presence of consciousness.

2. The role of what matters in constitution

The second difference between my account of personal identity vis-à-vis those of Snowdon, Olson and Baker is the following. Promoters of animalism and the constitution view alike take it for granted that what makes us us[14] (Olson 1997), or what makes a human person a human person[15] (Baker 2000), ought to be something that is itself fundamental and important to this being.[16]

For Baker, this is the first-person's perspective. As we saw, a first-person's perspective for Baker is not just a point of view. Animals that

---

[14] "When does one begin? I do not mean when human life begins, or whether an embryo or fetus is a human being. My question is when *you* and *I* began to exist" (Olson 1997, p. 90, my italics). Olson also doubts that any philosopher studying philosophy has meant to ask whether someone is still the same person. He says that even if they asked this, they actually meant to ask whether someone is still you (Ibid, p. 25).

[15] "'What am I' and 'What is a person?' – are distinct questions…..What I am most fundamentally is a person….A person…is a being with a first-person perspective. But the kind of person that I am is a human person, necessarily embodied" (Baker 2000, p. 6).

[16] "According to the Constitution View, what is ontologically *most important* about human persons is that they are persons. According to the Animalist View, by contrast what is ontologically *most important* about human persons is that they are animals" (Baker 2000, p. 147, my italics).

Or: "…on the Constitution View, our unique characteristics and *what we care deeply about depend squarely on what we fundamentally are*: persons" (Ibid, p. 163).

Again: "It is a signal virtue of the Constitution View that it directly connects what is most important to us and about us to what we most fundamentally are" (Ibid, p. 164).

are not at the same time human persons may have a point of view. A cat, for example, has the intuition that she is situated at a particular place. This makes her jump in the right direction and exactly as far as needed to catch a mouse. Since cats are not persons, a point of view cannot serve as a distinguishing characteristic of personhood. However, a first-person's perspective can serve as such a distinguishing characteristic – at least as far as Baker is concerned. She defines a first-person's perspective as the idea that someone can have of being a self. A first-person's perspective manifests itself when someone experiences herself as herself.[17] Two instances of the manifestation of this perspective are the following. (1) Someone may have the thought 'I look awful today.' (2) She may also have the thought 'I don't feel well today.' This person does not just see something awful, nor does she just feel nauseous. She has the additional thought that there is a subject with this look and feeling, and that this subject is not another person, but she herself. In *Persons and Bodies*, Baker (2000) repeatedly emphasizes that this first-person's perspective is fundamental in the constitution of human persons, because it is *important* for human persons.

> "However the first-person perspective came about, it is unique and unlike anything else in nature, and it makes possible much of what matters to us. It even makes possible our conceiving of things as mattering to us" (Baker 2000, p. 163, my italics).

---

[17] "So, although the dog has a first-person point of view, the person knows that she has a first-person point of view. She conceives of herself as the origin of a perspective and in doing so adds consciousness of unity to the unity of consciousness (that a dog has). Call the unity of consciousness underwritten by a first-person perspective 'strong unity of consciousness'" (Baker 2000, p. 162).

"The first-person perspective allows us, at least to some extent, to be self-conscious about our goals and to decide which ones to pursue and how to pursue them….animals have no control over their goals; but…we, unlike the rest of the animal kingdom, have a certain control over (some of) our goals…" (Ibid, p. 14).

Further, "[o]nly a being with a first-person perspective can have a grip on the fact that she even has a character or that she has habits at all. Only a being with a first-person perspective can evaluate herself, find herself wanting, and try to change herself in various ways" (Ibid, p. 15).

Again: "Let me …canvass a number of things…that are pretheoretically our *important* features and then show how they are made possible by the first-person perspective….

In the first group are cognitive abilities that we have only because we have first-person perspectives: (1)…we are the only beings who know that we are going to die…only a being with a first-person perspective could know that she – she, herself – was going to die eventually….(2) We can envisage many alternative possibilities for our own futures. To imagine oneself in this or that situation requires conceiving of oneself as oneself, in the first person, not as someone who fits a description or is picked out by a third-person demonstrative…(3) Only beings with first person perspective could make sense of the ancient dictum 'Know thyself'….

In the second group are practical abilities that we have only because we have first-person perspectives. (1) We can have life projects and plans; we can choose our ideals and assess our desires and try to change them to conform better to our ideals. One can try to rid oneself of a desire only if one can conceive of one's desire as one's own, and to conceive of one's desire as one's own is to have a first-person perspective. (2) We do not simply act in accordance with laws of nature. We also follow rules and even make up rules for ourselves….(3) We can decide what matters to us (within limits)….(4) We can ask, 'What am I? Who am I? What kind of life ought I to lead?'" (Ibid, pp. 159-160, my italics)

These kinds of observation are also what fuels Baker's criticism against animalism. She reproaches animalists for not taking persons seriously (Ibid, pp. 218-222).

Notwithstanding this critique, animalists equally rule that everything that constitutes a particular kind must be judged by us as important for this kind. The latter is a precondition for the former: according to the animalists, something can only constitute a particular kind, if we judge it to be important for this kind.

At first sight this may not seem to be the case. Olson, for instance, mentions that he has

> "tried to divorce numerical identity from those relations of practical concern that are traditionally thought to go along with our identity: whose actions we are accountable for, how we ought to be treated over time, and (perhaps) who deserves our prudential concern" (Olson 1997, p. 70).

He also says that

> "the consequences the Biological Approach has for ethics would be … surprising, for it would entail that those practical relations did not coincide reliably with psychological continuity. I could be responsible for something I did at an earlier time even if I am no longer psychologically continuous with myself as I was then. And I might fail to be responsible for actions done by someone who had my brain and my psychology, which I can remember and feel responsible for" (Ibid, p. 71).

At first sight then, Olson seems to distinguish what preserves numerical identity from what typically matters to us in this preservation: just as Parfit (1984), he points out that what we really care about when we seem to care about the survival of a person – i.e. the continuation of his projects, character, thoughts – does not itself

guarantee numerical identity, and he demonstrates that what seems to be important for the continuance of a person's numerical identity – i.e. psychological continuity – is, in fact, not always preserved when his numerical identity is preserved.

Still, Olson assumes that, what constitutes a kind, must be of intrinsic importance to this kind. As other animalists, Olson reasons that, since a certain organization between elements is clearly necessary for the maintenance of life, and we judge that the continuation of life is important for an organism, this life and this organization must be ontologically constitutive of an organism. This becomes clear when he explains why "you are the organism that began to exist roughly fourteen to seventeen days after conception" (Ibid, p. 91). According to Olson, this is so because this is "when the cells that develop into the fetus (as opposed to the placenta) become specialized and begin to grow and function in a coordinated manner" (Ibid, p. 91).

> "Until the end of the second week after fertilisation, the cells are all alike, or omnipotent: they do not have specialized tasks, and each can be the ancestor of any kind of human cell. Most of those cells in fact develop into the placenta and other supporting structures and not into the embryo proper. Each functions independently of the others, metabolizing and dividing at its own rate. If you separated the cells into two clumps, you would end up with identical twins; and if you put the separated cells back together – before they begin to specialize, at any rate – only one human being will result. In fact, the cells can be rearranged arbitrarily without affecting the eventual outcome" (Ibid, p. 90).

So, according to Olson, you are not a fertilized egg, because this egg is not yet by itself (i.e. intrinsically) important for you, as the fetus that may grow into a person. It does not make you breathe and metabolize. It does not make you into an individual. So, it does nothing

that is of intrinsic importance to you. At most, it creates your *external conditions* – such as the beginning of the placenta – through which you can later come into being.

Contrary to what is the case for Baker and Olson, my research aim is not to find those elements that are in themselves constitutive of personal identity. Instead, I examine which elements factually have a role in the constitution of our idea of what constitutes personal identity. Subsequently, I attempt to explain why we attribute such constituting role to these elements. This explanation is not necessarily a justification. We may turn out to treat elements as important for this constitution that are neither intrinsically important for the constitution of personal identity, nor for personhood, merely because they are typically close to something that is necessary to this constitution. We would then attribute it with constituting power through mere association.

3. The body as symbol

This brings me to a third difference between my treatment of the role of the body for the constitution of personal identity and those of Baker and Olson. Contrary to Baker and Olson, I stress how the body can become a symbol of the person and examine what the role of this symbol can be in the constitution of (our idea of) personal identity.

As far as the body is concerned, Baker and Olson are interested in whether, and how, natural properties of the body are properties of one animal or person. Baker (2000), for instance, says that there is not only a body that is heavy, but also a person who is derivatively heavy,

because her body is non-derivatively heavy. Olson claims that you survive as long as your brainstem and midbrain organs remain intact (Olson 1997, p. 116), for they regulate the life-sustaining functions of an organism (Ibid, p. 45), such as its breathing, circulation of blood and metabolism (Ibid, p. 16). So, according to Olson, you go where your whole brain goes (Ibid, p. 45).[18]

I recognize that certain natural characteristics of the body have a role in the constitution of (our idea of) what a person is and that there are such characteristics whose presence can make a person stay the same. Yet, in addition, I wish to highlight that the body, in part, gets its role in this constitution, because it becomes a symbol of the person.

Let me illustrate this difference in approach by referring to one of Snowdon's defenses against the claim that persons are something over and above animals. Snowdon concedes (Snowdon 1995, p. 82) that some people will, in certain cases, say that the person is departed while looking at her remaining living organism. This may happen when they visit a comatose patient in a vegetative state. Here, the person and the animal do not, at first sight, seem to coincide. However, Snowdon counters the conclusion to which this case seems to lead by quoting another aspect of the case. He asks us to imagine what we would think, if the vegetative state, in which the patient has sunk, would make the doctor tell us that the family member, whom we are looking for, is no longer in the hospital. Snowdon predicts that we would find this

[18] Note that Olson does not hold that you go where your cerebrum goes. According to Olson, you could survive without your cerebrum, as long as the part of the brain that controls your life-sustaining functions stays preserved. To hold that you follow your cerebrum is to assume a psychological criterion for your survival: the cerebrum makes your mental performances possible (Olson 1997, p. 18, 21, 44).

statement premature. He concludes that, after all, we still see the comatose patient as a person, i.e. as one of our family members.

I follow Snowdon thus far. However, contrary to Snowdon, I would not explain this case by saying that we still consider this comatose patient as our family member because he is still the same living organism. Instead, I would say that this artificially breathing organism retains the value of our family member because our family member has lived through it for such an extensive time. This organism does not leave us indifferent, because it has harboured our family member. We associate this organism so much with our family member that we would perceive a disrespectful treatment of this body as disrespectful behaviour towards this person. The organism gets the value of the person because we associate it with the person and because it becomes a symbol for this person, not because it is the person.

Olson is somewhat aware of the possibility that an organism could become a symbol of a person. This becomes clear when he wonders whether there can be a similarity between how an organism can, in some cases, receive its meaning or value and how a statue does:

> "My question in the Vegetable Case is whether the human vegetable that results when your cerebrum is destroyed is strictly and literally you, or whether it is no more you than a statue erected after your death would be you. Do you come to be a human vegetable, or do you cease to exist and get replaced by a vegetable, much as you might be replaced by a statue" (Olson 1997, p. 9)?

However, Olson denies that this is the case:

> "…there is another kind of continuity to be considered. Your life-sustaining functions are not disrupted when you lapse into a persistent vegetative state... Nor is this simply 'bodily' continuity. When you die, you may leave behind a corpse that preserves your gross anatomical structure and is made of the same matter as you

> were recently made of. But the human vegetable in our story is related to you in a far more interesting way than any corpse could be. The vegetable, but not the corpse, has inherited your biological life" (Ibid, p. 11).

If one would, contrary to Olson, not immediately choose for the ontological view on this case, and if one would not dismiss the symbolic view on this case, one could explain why some people will be more inclined than others to call a comatose patient in a vegetative state a person. There is no strict boundary that determines whether something will still receive the symbolic value of something else through its association with this something else. Also, the symbolic sensitivity of persons may differ. The clothes of a famous person may get value because he wore them. The same goes for the car with which he drove. But what with the road on which he drove? For some of his fans, this will not have a special value, for others it may.[19]

My regard for the symbolic value that an organism may get makes me differ from both Olson and Baker on yet another point.

Olson holds that "[t]here could not be non-people who were exactly like people but for their persistence conditions" (Olson 1997, p. 108). Baker similarly contends: "…on the Constitution View, it is impossible that something could be psychologically just like a person without really being a person" (Baker 2000, p. 197). I disagree. I hold

[19] This was also recognized by Kripke: "… my characterization has been far less specific than a real set of necessary and sufficient conditions for reference would be. Obviously the name is passed on from link to link. But of course not every sort of causal chain reaching from me to a certain man will do for me to make a reference" (1972, p. 93).

that our idea of people entails that all living creatures that are produced and born by people are people.

Baker explicitly denies this:

> "To suppose that any product of natural selection must be understandable in wholly biological terms is to commit what used to be called the 'genetic fallacy': the fallacy of assuming that where something came from determines what it is. I am interested here in what we are, not in where we come from" (Ibid, p. 18).

Yet, if a person gives birth to a severally disabled baby, we will still hold that this baby belongs to the class of people. If a chimpanzee gives birth to a very smart and human looking baby, we will call it an extraordinary chimpanzee. To determine whether babies belong to the class of people or not, we do not look at their characteristics, but at their origin. If the origin does not guarantee specific characteristics, this can only be so because we so often see people come forth from a specific origin, that we must associate them with such an origin. This origin then gets its meaning through association, not through characteristics it guarantees. The origin gets its people-constituting power because it belongs to a field of meaning – i.e. a symbolic field –, not because it generates specific ontological features.

With the indication that our body may have a role as symbol in the constitution of our (idea) of personhood and personal identity, I have pointed to the third of three differences between my account of personal identity and those of Olson, Snowdon, and Baker. By elaborating on these differences, I hope to both have made clear which actual roles of the body in the constitution of personal identity I will not consider, and have broadened the idea of the kinds of roles that the body could have in this constitution. In the next three chapters, I will

further analyze what can be the role in this constitution of the body, considered as material token, and examine how the nature of our imagination, and of material bodies, can makes us attribute such a role to the body.

CHAPTER 4:
MATERIAL BODIES AND PERSONS AS PRIMARY PARTICULARS

Now that I have determined my question as 'what makes us still consider someone as numerically the same?' (chapter 1) and have made plausible that this could be the continuation of one particular living organism that still generates consciousness (chapter 2) – indicating that this is not just the case because this organism preserves life, nor because it can produce consciousness, but also because it is materially continuous (chapter 3) –, I will examine which factors can help explain this role of our body in our recognition of a person's numerical identity.

I. MATERIAL BODIES FUNCTION AS PRIMARY PARTICULARS

For a first explanation of this role of the body, we can turn to the work of P.F. Strawson. In his *Individuals: An Essay in Descriptive Metaphysics* (1959), he demonstrates how, as far as re-identification is concerned, material bodies function as primary particulars in our conceptual scheme. My claim will be that, when material bodies generally help us to re-identify particulars, they can also help us to re-identify a particular person – even when we think of this person as being more than his material body alone. To see how this works, let's

first look at what Strawson attempted to discover and at his reasons for calling material bodies 'primary particulars' in our way of re-identifying particulars.

P.F. Strawson develops a descriptive metaphysics. In doing so, he intends to describe the actual structure of our thought.[20] He seeks to determine what our understanding is of the world we live in, and of which concepts, ways of identifying, referring and predicating we make use in this understanding. It is not his objective to determine what is absolutely logically necessary in order to think in a particular way. Our way of, and means for, understanding may be contingent, yet actual. Nor is Strawson interested in providing a genealogy of our structure of thought. He does not examine at which moment in time, and for what reason, we started to use a particular concept. Strawson judges that the subject of his research is too fundamental for this. He is interested in highlighting those concepts that are so fundamental for the human understanding of the world we live in, that they do not seem to be historical.[21] It would moreover be hard to imagine how a human understanding of the world would have looked like without our

---

[20] "Descriptive metaphysics is content to describe the actual structure of our thought about the world…." (Strawson 1959, p. 9)

[21] "The structure he seeks does not readily display itself on the surface of language, but lies submerged" (Strawson 1959, p. 10).

"…there is a massive central core of human thinking which has no history – or none recorded in histories of thought; there are categories and concepts which, in their most fundamental character, change not at all…They are the commonplaces of the least refined thinking…It is with these, their interconnexions, and the structure that they form, that a descriptive metaphysics will be primarily concerned" (Strawson 1959, p. 10).

grasp and usage of these concepts. We may try to imagine such a scenario, but then we would again be thinking about a potential structure of thought that is not our current actual structure, and this is not what Strawson sets out to examine.

Strawson starts from the observation that we live in a spatio-temporal world: it is a fact that we perceive everything as happening at one particular time in space. He also observes that we frequently identify and re-identify particulars. We say of things and creatures that they are particular creatures and, when we see them again or speak about them, we identify them as those particulars that we encountered before. Because this identifying is so prominent in our thinking and speaking, and is what first allows us to live in a *common* world where we can identify and refer to the same things, Strawson decides to examine how we identify particulars. He concludes that material bodies are primary particulars in this process: they are particulars to which we need to refer in order to identify other particulars, while we do not need to refer to other particulars to identify them.

Strawson concludes that material bodies are such primary particulars, after establishing that *descriptions* do not uniquely refer and hence cannot uniquely identify, or, in other words, pinpoint one specific particular. The possibility of massive reduplication prevents descriptions from doing so. Even when a description describes an object in incredible detail and mentions a list of relations between the described object and other objects, the possibility remains that somewhere in the world there is a second object which has all the

details mentioned and maintains completely similar relations with other objects.[22]

This problem of identification, or unique reference, is solved when our reference is indicative or demonstrative, instead of merely descriptive. In an indicative reference, you refer to something by pointing it out, just as you would point something out with your index finger. You show or demonstrate where something is. As we can see in our use of demonstratives, we need a point of reference to indicate something. We know what a speaker means with 'this' or 'that', 'these' or 'those', because we can locate the speaker, and so know which things are closer or further away from him. Here, the speaker functions as a point of reference from where we can follow his indication. An indicative reference from such point of reference always uniquely identifies. It cannot accidentally refer to two similar particulars of the same kind[23] that find themselves within a similarly organized set of other particulars. For, however similar the patterns of particulars around two resembling particulars are, the spatial world, in which we

[22] "However much one adds to the description of the sector one knows about – its internal detail and its external relations – this possibility of massive reduplication remains open. No extension of one's knowledge of the world can eliminate this possibility. So, however extensive the speaker's knowledge and however extensive the hearer's, neither can know that the former's identifying description in fact applies uniquely" (Strawson 1959, p. 20).

[23] I add 'of the same kind', because it is possible that we stand in the same spatial relationship to two particulars of a different kind. There can, for instance, be the same spatial relationship between me and a particular person, and between me and this particular person's body. This body and this person are then two different particulars (even though they have some properties in common, they are of a different kind), which nevertheless find themselves at the same place.

live, does not allow that the respective spatial relations between a point of reference and two distinct particulars are the same.[24]

When we identify or refer to particulars, we, of course, do not always use demonstratives, such as 'this' or 'that'. We also use proper names, such as (a) 'the U.S.A.' or (b) 'Charlotte'; (c) we use the calendar time to refer to a particular event; and (d) we often deduce who or what someone is talking about from the context of a story. Consequentially, we do not always notice that our identification of particulars is demonstrative in kind. Still, all our identifications ultimately depend on a demonstrative identification. (a) When we use a proper name to refer to a place, we only understand where this place is, when we understand where it is with regard to us, the reference point and point of orientation. The fact that Spain is to the south of France only means something to me, when I know what it would mean for me to move south. (b) Our reference to a person by means of a proper name equally counts as a demonstrative indication. A person's name is not used as an abbreviated description of a particular character. I won't stop calling a particular person 'Charlotte' when she evolves from being a difficult learner into a successful business woman. Instead, I use the name 'Charlotte' to refer to one particular person who came into

[24] "For all particulars in space and time, it is not only plausible to claim, it is necessary to admit, that there is just such a system: the system of spatial and temporal relations, in which every particular is uniquely related to every other. The universe might be repetitive in various ways. But this fact is no obstacle in principle to supplying descriptions of the kind required. For by demonstrative identification we can determine a common reference point and common axes of spatial direction; and with these at our disposal we have also the theoretical possibility of a description of every other particular in space and time as uniquely related to our reference point. Perhaps not all particulars are in both time and space. But it is at least plausible to assume that every particular which is not, is uniquely related in some other way to one which is" (Strawson 1959, pp. 22-23).

existence at a certain moment, has since then always been at just one place at a time, and followed a continuous, i.e. non-interrupted, track through space.[25] She never disappeared at one place to then pop up at a totally different place. (c) We can only meaningfully use calendar time to refer to a particular event, when we know how the time, now, relates to that time. There is only a 'now' for us – beings who can serve as their own point of reference. No notion of 'now' could be found in an unconscious system. (d) It is true that the context of a story often makes it clear to us, to whom the storyteller is referring. However, this does not mean that the reference to this person can depend on descriptions alone. It still involves some demonstrative indications. When someone says that he had dinner 'with him', the context of the story may make it obvious to me that this 'him' is his friend. Still, my understanding of who this friend is, does not only follow from the way in which this friend was once described to me, it also involves my understanding that this friend is one particular being who, as long as he exists, always finds himself at one particular place in space and can be encountered there by me or others. (e) Even supposedly 'pure individuating descriptions,' which start with phrases like 'the only…' or 'the first…,' do not refer to unique particulars all by themselves. They can only do so in combination with a demonstrative indication.[26]

---

[25] This was already remarked by Kripke (1972, lecture 1).

[26] "Descriptions can be framed which begin with phrases like 'the only…' or 'the first…' and thus proclaim as it were, the uniqueness of their application. Let us call them 'logically individuating descriptions'" (Strawson 1959, p. 26). "A pure individuating description, like any other logically individuating description, may fail of application not only when there are no candidates for the title, but also when there are two or more candidates with equally good and hence mutually destructive claims

For, a supposedly pure individuating description would fail to identify a unique individual, if there either is no individual, or if there are two individuals to which this description applies. This risk can only be avoided, if one knows, for a fact, that only one such individual exists, and this knowledge can only be provided by one's exploration of the world. Since the latter requires that we start from, and take ourselves as, a reference point, so as to be able to determine what we did and did not yet explore, the supposedly pure individuating description never individuates by its mere description. And "[e]ven if it were possible to satisfy the formal conditions of particular-identification in a way which left the particular completely detached and cut off, as it were from the general unified framework of knowledge of particulars, the achievement would be a peculiarly useless one. So long as our knowledge of it retained this completely detached character, the particular would have no part to play in our general scheme of knowledge; we could for example, learn nothing new about it except by learning new general truths" (Strawson 1959, p. 28).

The previous should have made clear that "the system of spatio-temporal relations has a peculiar comprehensiveness and pervasiveness" (Ibid, p. 25), which, according to Strawson, qualifies "it uniquely to serve as the framework within which we can organize our

---

and no candidate with a better claim… The only safe way, in general, to elaborate the description sufficiently to eliminate the one risk, without increasing the other, would be to draw on our actual knowledge of stretches of the world and its history; but in so far as we do this, we can no longer sincerely claim to be unable to connect our description at any point with items belonging to the unified framework of our knowledge of particulars" (Ibid, pp. 27-28).

individuating thought about particulars. Every particular either has its place in this system, or is of a kind the members of which cannot in general be identified except by reference to particulars of other kinds which have their place in it" (Ibid, p. 25).

Strawson then concludes that it is our spatio-temporal framework that allows us to live in a common world and identify and refer to the same things: "By means of identifying references, we fit other people's reports and stories, along with our own, into the single story about empirical reality; and this fitting together this connexion, rests ultimately on relating the particulars which figure in the stories in the single spatio-temporal system which we ourselves occupy" (Ibid, p. 29). Strawson grants that it may be possible for us to refer differently to particulars in another world, where our experiences would be different. He even explores if we could identify particulars in a world where there are only sounds (Ibid, pp. 59-86). Still, the bottom line for Strawson remains that it is not contingent, in our world, to identify particulars by situating them in a spatio-temporal framework. We perceive it as necessary or guaranteed that every particular has its place in this framework:

> "Suppose someone told of a thing of a certain kind, and of certain things that had happened to it; and, when asked where that thing had been, and when the events he recounted had occurred, said, not that he did not know, but that they did not belong at all to our spatio-temporal system, that they did not take place at any distance from here or at any distance of time from now. Then we should say and take him to be saying, that the events in question had not really occurred, that the thing in question did not really exist. In saying this we should show how we operate with the concept of reality… We are dealing here with something that conditions our whole way of talking and thinking, and it is for this reason that we feel it to be non-contingent" (Ibid, p. 29).

We do not only exploit our common spatio-temporal understanding of the world and history we live in, to identify, or point to, particulars. We also use it to *re*-identify them. Places are necessary to re-identify particulars. This must be so, to compensate for our interrupted observations. I can now point to a particular and say 'this is X', but I cannot count on my uninterrupted observation of X to re-identify it as the same X at a later time. I move and sleep, and so my observation of X gets interrupted. Still, I will oftentimes re-identify X by re-identifying the place at which I last saw it. Children, for instance, re-identify their own handbook by assuming that the one laying at the same place where they last saw it, is theirs. Or, if it is no longer there, they imagine that someone must have moved it from the one place, to another place. They then suppose that it was constantly somewhere in adjacent places, not that it dropped out of existence at one place, and reappeared at another. To be able to re-identify places like these, we must, in turn, be able to re-identify physical bodies, for "… places are defined only by the relations of things…" (Ibid, p. 37). If "we were never willing to ascribe particular-identity …[t]hen we should, as it were, have the idea of a new, a different, spatial system for each new continuous stretch of observation. … Each new system would be wholly independent of every other" (Ibid, p. 33).

A skeptic could point out that, if we re-identify particulars as the same when our observation of them is not continuous, we can never really be sure that they are in fact the same. Perhaps the object at time 2, which looks like an object at time 1, finds itself at a place that looks like the place where the qualitatively identical object at time 1 found itself, but still is not the same object – it might just be a similarly looking object at a similarly looking place, but, in fact, be an object at

a different place. This is true. This possibility exists. Only, it does not show that there is a possibility for us to identify objects in a different, perhaps more secure way, which does not appeal to our common understanding and usage of a spatio-temporal framework to identify places and objects. For, as Strawson wittingly points out, even the skeptic does not escape from using this framework to doubt whether we always re-identify correctly. If we did not assume one spatio-temporal framework, in which all particulars of the world have a place,

> "[t]here would be no question of doubt about the identity of an item in one system with an item in another. For such a doubt makes sense only if the two systems are not independent, if they are parts, in some way related, of a single system which includes them both….This gives us a more profound characterization of the skeptic's position. He pretends to accept a conceptual scheme, but at the same time quietly rejects one of the conditions of its employment" (Ibid, p. 35).

All of this makes clear that the spatio-temporal framework, in which we situate particulars plays a crucial role in particular identification. It is only of particulars, which could be situated in our spatio-temporal framework, that we are willing to say that they exist or have existed. Further, we can only uniquely identify those particulars that can be situated in such a framework. This unique identification, or reference, is ultimately indicative. Lastly, our *re*-identification of particulars, depends on the spatio-temporal framework, in which we situate them. We re-identify them by determining both the place at which they first found themselves and the path along which they traveled.

To this, Strawson adds that physical bodies function as basic particulars in our particular identification. They are the particulars to

which we ultimately need to refer, if we identify other particulars, and which do not, themselves, require a reference to another kind of particular to be identified.[27] Strawson does not provide any logically sufficient argument that proves that only material bodies are able to take up this role as basic particulars. Instead, he illustrates which assets of material bodies – in contrast with properties of other particulars – make them the best particulars to function as basic particulars. Material bodies get the latter function because their (a) observability,[28] (b) inter-individual diversity and "relatively fixed or regularly changing spatial relations" (Ibid, p. 53), and (c) intra-individual stability and endurance help us identify other particulars in reference to them, as well as make us identify places and so set up our spatio-temporal framework in which we can re-identify particulars.[29] (a) In the *common* framework, which we need to refer to the same things and talk about them, "[t]he minimum conditions of independent identifiability for a type of particulars [is] … that its members should be neither private nor unobservable" (Ibid, p. 53). With regard to this, three-dimensional material bodies have a clear

[27] Cfr. Strawson 1959, pp. 38-39.

[28] "The minimum conditions of independent identifiability for a type of particulars were that its members should be neither private nor unobservable" (Strawson 1959, p. 53).

[29] "[These particulars] must be three-dimensional objects with some endurance through time. They must be accessible to such means of observation as we have; and, since those means are strictly limited in power they must collectively have enough diversity, richness, stability and endurance to make possible and natural just that conception of a single unitary framework which we possess. Of the categories of objects which we recognize, only those satisfy these requirements which are, or possess material bodies – in a broad sense of the expression. Material bodies constitute the framework. Hence, given a certain general feature of the conceptual scheme we possess, and given the character of the available major categories, things which are, or possess material bodies must be the basic particulars" (Strawson 1959, p. 39).

advantage over and against sensations (to which we can only uniquely refer, when we attribute them to a particular person) and abstract theoretical constructs (such as a strike, which we can only identify by referring to particular men, tools and factories[30]). (c) In contrast with events and processes, material bodies typically endure longer. This is why, when we want to refer to an event or process that we do not witness at the moment, we often refer "to a place at which it was audible or visible [and which we identify by objects or geographical features carved into material bodies[31]], or to a particular material object which was causally connected with it" (Ibid, p. 47). In short: material bodies have the necessary assets to give us the means to construct the spatio-temporal framework in which we find ourselves[32,33] and to situate all particulars within this framework.[34]

Again, it is true that it is not always immediately obvious that our identification of particulars ultimately depends on our identification of material objects. We do not always refer by naming such objects. Still, our frequent use of proper names hint at which role material bodies play in our particular-identification: "among particulars, the bearers par excellence of proper names are persons and places. It is a

---

[30] Cfr. Strawson 1959, p. 44.

[31] Cfr. Strawson 1959, p. 53.

[32] "Given a certain general feature of the conceptual scheme we possess, and given the character of the available major categories, things which are, or possess material bodies must be the basic particulars" (Strawson 1959., p. 39).

[33] "The fact that identification in general has a temporal as well as a spatial aspect is no objection. For material bodies, or things which have them exhibit relations between themselves which have a temporal aspect. One thing replaces or begets another. Things pass through places" (Strawson 1959, p. 54).

[34] "Material bodies, in a broad sense of the word secure to us one single common and continuously extendable framework of reference, any constituent of which can be identifyingly referred to without reference to any particular of any other type" (Strawson 1959, p. 54).

conceptual truth, as we have seen that places are defined by the relations of material bodies; and it is also a conceptual truth…that persons have material bodies" (Ibid, p. 58).

All of this together then, gives us a first explanation of what makes us identify and re-identify persons through their bodies, even when we consider mental characteristics to be necessary and typifying for persons. By giving us a location and orientation, the body can make us into a reference point that allows us to identify both ourselves and other particulars by their place and our own, and to re-identify them and ourselves by our respective plausible movements. As I mentioned in the second chapter, our material bodies are easier to identify than our thoughts and mental characteristics because material bodies are always publicly observable. They are easier to *re*-identify because they are always somewhere for as long as they exist. This means that, when our observation of these bodies was interrupted, we can still retrieve a particular body at the same place as we encountered it before or at a place which is thus located, that we can imagine that this body has moved there by now, by following a continuous track.

It is true that there is always the possibility that we will misidentify a particular body. When our observation of a body was interrupted, we may reasonably expect that a certain body is the same as the one that we saw before, just because we now see it at a place close to where we saw it before. Still, it could be another qualitatively similar body. In spite of this possibility of misidentification, the distinction between material bodies and mental characters remains such that the re-identification of the former is primordial, whereas the re-identification of the latter is not. We know what the conditions are for a material

body to count as the same numerical body we saw before: all its parts must either remain the same or have evolved from one another, and the body must have traveled through space along a continuous path. We do not have this with a mental character or someone's personal set of mental thoughts. If we do not identify this character or particular set of thoughts by means of referring to another particular (such as the person who has them), then we cannot re-identify a character or set of thoughts as numerically the same. Character traits of a particular personality and thoughts of a particular person do not all evolve from one another: a person can have two very different thoughts without the one having sparked the other; he can also establish two different non-connected ways of behaving (e.g.: he can be talkative and a hard worker). Further, character traits are not continuously exhibited and thoughts are not uninterruptedly had. Together, this means that we can neither re-identify character traits or a set of thoughts as numerically the same by finding a causal or logical relation between them, nor by just following them through time. This is not just so because our observation is interrupted. It is impossible because the nature of a character or set of thoughts does not allow for it. The fact that there is no clear alternative for re-identifying a character or personal set of thoughts, apart from recognizing it by recognizing the person (another particular who has them), means that only material bodies (which can be re-identified by appealing to bodily characteristics only) are the particulars that function as basic particulars in our identification and re-identification of persons. This can, of course, only be so, if a body is had by the same subject (or person), who can have the mental characteristics which we fail to re-

identify all by themselves. In the paragraph that now follows, I say more about how Strawson establishes that this is indeed the case.

II. We ascribe characteristics to a person, not to a body and mind, which only subsequently constitute a person

The main aim of this chapter is to explain why we oftentimes identify persons by their bodies, even when we judge that these bodies, as such, do not yet make them persons, as well as that a particular body is not what most typifies a particular personality. I just explained, with Strawson, that one of the reasons for this phenomenon is that bodies function as basic particulars in our way of identifying other particulars. Regardless of what we find important in other particulars, our identification of them ultimately depends on our identification of material bodies, which are basic particulars.

Strawson analyzes another phenomenon that renders our way of identifying, in these cases, more intelligible. If mental characteristics belonged to a mental subject, while physical characteristics belonged to a distinct subject that could be described as a purely material body, then it would be illogical to identify someone who has mental characteristics through an entity that has no mental characteristics. Strawson shows that identifying a person through his body is not illogical in this regard, because persons are no assembly of a respectively mental and material subject. He contends that when we analyze our structure of thought and speaking about persons, we will

find that we do not think of them as embodied minds or animated bodies[35]. Instead, our concept of a person is 'primitive'.

> "What I mean by the concept of a person is the concept of a type of entity such that both predicates ascribing states of consciousness and predicates ascribing corporeal characteristics, a physical situation &c. are equally applicable to a single individual of that single type….What I mean by saying that this concept is primitive…[is] that a necessary condition of states of consciousness being ascribed at all is that they should be ascribed to the very same things as certain corporeal characteristics, a certain physical situation &c" (Strawson 1959, pp.101-102).

So, in the case of a person, we do not ascribe mental characteristics to a mental subject and physical ones to a bodily subject. We ascribe both characteristics to one and the same subject, i.e. a particular person.

> "We ascribe to ourselves actions and intentions (I am doing, did, shall do this); sensations (I am warm, in pain); thoughts and feelings (I think, wonder, want this, am angry disappointed, contended); perceptions and memories (I see this, hear the other, remember that). We ascribe to ourselves, in two senses, position: location (I am on the sofa) and attitude (I am lying down). And of course we ascribe to ourselves not only temporary conditions, states, situations like these but also relatively enduring characteristics, including physical characteristics like height, colouring, shape and weight" (Ibid, p. 89).

Strawson admits that it is tempting to explain this phenomenon by referring to the special causal role of our body in our experiencing. Many of our experiences, of which we are aware or mindful, do indeed depend on, and are influenced by, the kind of body we have and its positioning. I only see something when my eyelids are open, when

[35] Cfr. Strawson 1959, p. 103.

what I see is in my field of vision, and when, upon this particular something, my vision is focused (which is due to the positioning of my pupils, the shape of my eyeballs, and my overall mental state). One could then think that our body connects whom we are as a mental being with whom we are as a physical being, and so gives us the idea of being a psycho-physical entity. Yet, Strawson points out that this special causal role of a particular body in our experiencing at most explains why we are attached to a *particular* body or determine that if one body is ours, this must be it. It does not explain where our idea of being a self or person comes from, to which we subsequently ascribe predicates, and of whom we subsequently think that it has a body.[36]

If not because of this special role of the body, why then is the concept of the person primitive? Why do we attribute experiences to a self, and why do we attribute physical and mental predicates to the same self? Strawson argues that this is necessary in our conceptual scheme. This becomes clear when thinking otherwise appears to be incoherent. A philosopher could try to contest the primitiveness of a self or person and say that there is no such thing as a person who keeps particular sets of experiences together. Yet,

---

[36] "We may summarize such facts by saying that for each person there is one body which occupies a certain causal position in relation to that person's perceptual experience, a causal position which in various ways is unique in relation to each of the various kinds of perceptual experience he has…" (Strawson 1959, p. 92). "They explain …why I feel peculiarly attached to what in fact I call my own body; they even might be said to explain why, granted that I am going to speak of one body as mine, I should speak of this body as mine. But they do not explain why I should have the concept of myself at all, why I should ascribe my thoughts and experiences to anything…. They do not explain the concept we have of a person" (Ibid, pp. 93-94).

> "[w]hen he tries to state the contingent fact, which he thinks gives rise to the illusion of the 'ego', he has to state it in some such form as 'All *my* experiences are …uniquely dependent on the state of…body B'. For … [t]he proposition that *all* experiences are causally dependent on the state of a single body B…is just false" (Ibid, pp. 96-97, my italics).

Hence,

> "this account of the matter is not coherent…in that one who holds it is forced to make use of that sense of possession of which he denies the existence..." (Ibid, p. 96). "He must mean to be speaking of some class of experiences of the members of which it is in fact contingently true that they are all dependent on body B. The defining characteristic of this class is in fact that they are 'my experiences' or 'the experiences of some person', where the idea of possession expressed by 'my' and 'of' is the one he calls into question" (Ibid, p. 97). "States, or experiences, one might say, owe their identity as particulars to the identity of the person whose states or experiences they are" (Ibid, p. 97). "[T]he theorist could maintain his position only by denying that we could ever refer to particular states or experiences at all; and this position is ridiculous" (Ibid, p. 98).

After establishing that experiences are always ascribed to a person, Strawson contends that I can only ascribe a predicate to myself as such a person, when I am also prepared to ascribe these predicates to a person who is not me. If the latter were not the case, then I would not be able to ascribe predicates to myself as a person either.[37] Strawson does not explicitly explain why he is convinced that this is the case,

[37] "There would be no question of ascribing one's own states of consciousness, or experiences, to anything, unless one also ascribed, or were ready and able to ascribe, states of consciousness, or experiences, to other individual entities of the same logical type as that thing to which one ascribes one's own states of consciousness. The condition of reckoning oneself as a subject of such predicates is that one should also reckon others as subject of such predicates" (Strawson 1959, p. 104).

but it seems likely that he reasons (1) in a Wittgensteinian manner that we must be following some kind of rule regarding the ascription of predicates to a subject, i.e. that we only ascribe predicates to a subject, when we judge on certain grounds that something is the case. To this, he may add that this rule, and the concept of 'predicates' and 'subjects', only have meaning if we know when and how we can apply this rule again, and if we could re-identify other subjects.[38] (2) He may reason that there cannot just be one subject in the world for us, because we would then not be able to individuate this subject from others. If there is no individuation, then there is no subject or individual.[39]

Once he established that I ascribe a particular set of predicates, i.e. mine, to a particular subject, i.e. me, and that we also recognize others as such subject, Strawson describes how we conceive of ourselves and others as similar subjects, even though our way of ascribing predicates to either one is different: we ascribe predicates to ourselves without external observation of ourselves, but we only ascribe particular

[38] "… one should be able to distinguish from one another, to pick out or identify, different subjects of such predicates, i.e. different individuals of the type concerned" (Strawson 1959, p. 104).

[39] "It sometimes happens, with groups of human beings, that, as we say, their members think, feel and act 'as one'. I suggest it is a condition for the existence of the concept of an individual person, that this should happen only sometimes." (Strawson 1959, p. 114) "It is quite useless to say, at this point: 'But all the same, even if it happened all the time, every member of the group would have an individual consciousness, would embody an individual subject of experience.' For, once more, there is no sense in speaking of the individual consciousness just as such of the individual subject of experience just as such, there is no way of identifying such pure entities…as soon as we adopt the latter way of thinking, then we abandon the former" (Ibid, p. 115).

predicates to others after we observed them. Strawson uses this phenomenon to illustrate that we think of human subjects as psycho-physical beings, and not just as essentially mind, or merely embodied minds.[40] When we would say, of ourselves, that we are depressed, we would not do this on the basis of outer observation: we know immediately that we do not have energy, do not want to get out of bed, and feel burdened with fears. When we identify someone else as depressed, we do this on the basis of observation. We see that he is low in energy, scared, and does not want to get out of bed. What we see is not just a sign of his invisible depression. They are symptoms of this depression, i.e. they are elements that are part and parcel of what we

[40] "One can ascribe states of consciousness to oneself only if one can ascribe them to others. One can ascribe them to others only if one can identify other subjects of experience. And one cannot identify others if one can identify them only as subjects of experience, possessors of states of consciousness" (Strawson 1959, p. 99-100). "The condition, in turn, of this being possible is that the individuals concerned, including oneself, should be of a certain unique type of a type, namely, such that to each individual of that type there must be ascribed, or ascribable, both states of consciousness and corporeal characteristics…" (Ibid, p. 105).

call depression.[41] Apparently, the depression has characteristics that are both sensible and observable, both involving mind and body.[42]

---

[41] "… one ascribes P-predicates to others on the strength of observation of their behavior;…the behavior-criteria one goes on are not just signs of the presence of what is meant by the P-predicate, but are criteria of logically adequate kind for the ascription of the P-predicate… The point is not that we must accept this conclusion in order to avoid skepticism, but that we must accept it in order to explain the existence of the conceptual scheme in terms of which the skeptical problem is stated" (Strawson 1959., p. 106).

[42] "…the ascribing phrases are used in just the same sense when the subject is another as when the subject is oneself…" (Strawson 1959, p. 99) "…there is a kind of predicate which is unambiguously and adequately ascribable both on the basis of observation of the subject of the predicate and not on this basis, i.e. independently of observation of the subject the second case is the case where the ascriber is also the subject" (Ibid, p. 108).

e.g.: "in order for there to be such a concept as that of X's depression, the depression which X has, the concept must cover both what is felt, but not observed, by X, and what may be observed but not felt, by others that X (for all values of X). But it is perhaps better to say: X's depression is something one and the same thing, which is felt, but not observed, by X and observed, but not felt, by others that X….To refuse to accept this is to refuse to accept the structure of the language in which we talk about depression" (Ibid, p. 109).

"When we take the self-ascriptive aspect of the use of some P-predicates, say 'depressed', as primary, then a logical gap seems to open between the criteria on the strength of which we say that another is depressed, and the actual state of being depressed. What we do not realize is that if this logical gap is allowed to open, then it swallows not only his depression, but our depression as well. For if the logical gap exists, then depressed behavior, however much there is of it, is no more than a sign of depression. But it can only become a sign of depression because of an observed correlation between it and depression. But whose depression? Only mine, one is tempted to say. But if only mine, then not mine at all" (Ibid, p. 109).

We can now summarize Strawson's reasoning as follows. If we read Strawson with the main question of this chapter in mind, then we see how he, through an analysis of our conceptual scheme, explains why we can identify persons by means of their bodies, even when we think that persons have minds and that these are what typifies them most and what is most important to them. We do so because, in our world, which we experience as spatio-temporal, all individuation and re-identification of particulars ultimately depends on the possibility of re-identification of particular bodies. We can do so because we do not think of persons as assemblies of two different subjects (the mind and the body) but, instead, conceive of persons as subjects to whom physical, psychological, and psycho-physical predicates can be ascribed. This means that we can know where a person, as a complete

---

"we have to do with a class of predicates to the meaning of which it is essential that they should be both self-ascribable and other-ascribable to the same individual, where self-ascriptions are not made on the observational basis on which other-ascriptions are made but on another basis. It is not that these predicates have two kinds of meaning. Rather, it is essential to the single kind of meaning that they do have, that both ways of ascribing them should be perfectly in order" (Ibid, p. 110).

"If one is playing a game of cards, the distinctive markings of a certain card constitute a logically adequate criterion for calling it, say the Queen of Hearts; but in calling it this, in the context of the game, one is ascribing to it properties over and above the possession of these markings. The predicate gets its meaning from the whole structure of the game. So with the language in which we ascribe P-predicates. To say that the criteria on the strength of which we ascribe P-predicates to others are of a logically adequate kind for this ascription, is not to say that all there is to the ascriptive meaning of these predicates is these criteria. To say this is to forget that they are P-predicates, to forget the rest of the language-structure to which they belong" (Ibid, p. 110).

psycho-physical being, is, when we know where his body is. Strawson came to the conclusion that we conceive of persons as psycho-physical beings by discovering that we self-ascribe some predicates, and by inferring that we must ascribe these predicates to ourselves as a subject; that if we have a concept of ourselves as a subject, we must also think of others as subject; and that given that these others must be of the same logical type as we are, although we nevertheless ascribe predicates to them after external observation only, while we would self-ascribe them after merely feeling them, there must be personal predicates which can be both felt and externally observed. It is part of our conceptual scheme and language system that we can observe that a personal predicate applies either by feeling it, or by observing it in someone else.

After this conceptual explanation, Strawson offers us some examples of phenomena that we describe with predicates referring both to intention and bodily movement. Strawson thinks of "such things as 'going for a walk', 'coiling a rope', 'playing ball', 'writing a letter'"(Ibid, p. 111). We ascribe these predicates to others through observation, but exactly because these predicates refer to a substantial bodily movement, we are not tempted to say that this movement is just an indication of a mental subject who undergoes this. "[W]e see such movements as *actions* …" (Ibid, p. 112). Such natural phenomenon makes … it seem intelligible to us… that we have the conceptual scheme we have" (Ibid, p. 112). If there is observable behaviour that is

clearly intentional, then it is less strange that we use the same predicates for observable and intentional behaviour.[43]

The fact that our concept of a person is primitive, and not just an assembly of the separate subjects 'body' and 'mind', does not prevent us from having a notion of a corpse (i.e. a soulless human body) or of a disembodied person. However, these two latter notions still depend on the primitive concept we have of the person. Our idea of a corpse is an idea of a person who *no longer* has a soul. To think of one's

> "survival of bodily death…[o]ne has simply to think of oneself as having thoughts and memories as at present, visual and auditory experiences largely as at present, even, perhaps – though this involves certain complications – some quasi-tactual and organic sensations as at present, whilst (a) having no perceptions of a body related to one's experience as one's own body is, and (b) having no power of initiating changes in the physical condition of the

[43] We ascribe many of these predicates, which indicate that we do not conceive of the body as if it is as external to a person as his clothes are. We do this by developing names for the (partly) bodily behaviour of persons that distinguish it from similar behaviour of unconscious bodies. For instance, we distinguish the raising of an arm from a spasm; kneeling from bending; and screaming from making noise. 'Sleeping', as well is a predicate that is attributed to a person, and not just to a body or mind. Sleeping is something that takes place over time. It is a specific physical state of an organism that has some duration. At the same time, it is a state that is only seen as this specific state, in reference to conscious states that must be had by the same subject in the past and in the future. If the being asleep now was not conscious in the past, it would just be a non-conscious being, not a sleeping being. If this sleeping being now would not become conscious in the future, it would not be asleep, but in an irreversible coma. The same holds for many other phenomena, such as raising one's arm. It takes some time for a being to raise its arm, and it can only raise its arm, if it had the intent to raise its arm beforehand. If not, there would not be the raising of an arm, but merely a spasm.

> world….Condition (a) must be expanded by adding that no one else exhibits reactions indicating that he perceives a body at the point which one's body would be occupying if one were seeing and hearing in an embodied state from the point from which one is seeing and hearing in a disembodied state…" (Ibid, p. 116).

However, there are two implications in the latter case.

> "The first is that the strictly disembodied individual is strictly solitary…The other, and less commonly noticed point, is that in order to retain his idea of himself as an individual, he must always think of himself as *dis*embodied, as a *former* person….Since …he has …no personal life of his own to lead, he must live much in the memories of the personal life he did lead; or he might…achieve some kind of attenuated vicarious personal existence by taking a certain kind of interest in the human affairs of which he is a mute and invisible witness – much like that kind of spectator at a play who says to himself: 'That's what I should have done (or said)'" (Ibid, p. 116).

The primitiveness of the person, with regard to the bodily, psychological, and psycho-physical predicates that can be ascribed to them, remains. So does the function of material bodies, as basic particulars, in our reidentification of particulars in general, and of persons in an application of this phenomenon. By elaborating on this finding, I have given a first explanation of why we, oftentimes, identify persons by means of their still numerical identical body. Bodies are basic particulars which allow us to identify other particulars, such as persons. And, given that a person is not a composite of a mental and physical subject, but one particular to which mental, physical and psycho-physical characteristics are ascribed, we can identify a being who has psychological characteristics through another of his constitutive elements, i.e. the body. In the two following chapters, I will provide further illustrations and explanations of why the material body which, at first sight, says nothing about a person's personality,

can still have crucial role in the constitution of a person's numerical identity.

CHAPTER 5:
SELF-CONSCIOUSNESS'S SOURCE IS NOT JUST INTERNAL

I. INNER CONSCIOUSNESS IS NOT THE SOLE SOURCE OF SELF-CONSCIOUSNESS

In the previous chapter, I let Strawson point out that we can neither identify, nor re-identify a particular mind without ultimately identifying a body, as well as a person to whom both this mind and body can be attributed: our identification of any particular depends on our identification and re-identification of material bodies, which function as basic particulars; and, in identifying a body, we can simultaneously identify a conscious person, because persons are subjects who have both a body and a mind and to whom physical, psycho-physical, and psychological characteristics can be ascribed. An opponent of this view may admit that we can only identify *other* conscious beings as particular conscious beings by means of re-identifying their bodies, while he may simultaneously deny that the latter has any part in the constitution of the consciousness which I have of being a particular conscious being *myself*. He could argue that, since we do not have access to other minds, we can do nothing but rely on the assumption that particular conscious beings stay housed in respective particular bodies, and then identify these conscious beings through these bodies. At the same time, he could reason that we do have immediate access to our own minds, i.e. that our own mind

appears to provide full access to itself, and that it is consequently unnecessary for us to re-identify ourselves by means of re-identifying our body. I know that *I* feel something now, simply because I am feeling it, and not because I see that something happens to my body. Just like this, I can know that I was the one who did something in the past, merely because I remember it, and not because there is evidence that I was physically present at that event, at that time.

I argue against the idea that self-consciousness is a pure consciousness of consciousness at which one only arrives by consulting one's own consciousness. In this and the next chapter, I will demonstrate that many instances of our supposedly purely internal awareness of ourselves are in fact informed by something external to mere inner consciousness. This is the case for our experiential memory, as I will show in this chapter, and for our sensations and I-thoughts, which I will discuss in the next chapter. It should become clear that what we often take to be consciousness of just our minds, only comes about because we also have the idea that we take up space in the world and that others can follow us throughout this space. Once we learn that self-consciousness and, so ultimately, selves are constituted thus, we will have another answer to the question previously posed. We will better understand why we can identify ourselves (whom we take to be conscious beings) by means of the body which these selves may just seem to inhabit. It will have become clear to us that becoming aware of and, a fortiori, tracing a supposedly pure mental self, involve a reference to something physical.

II. PRELIMINARY TERMINOLOGY: SELF-CONSCIOUSNESS VS. SELF-KNOWLEDGE

I will make use of the terms 'self-consciousness' or 'self-awareness' on the one hand, and 'self-knowledge' on the other. I will define self-consciousness as a consciousness of ourselves which has, as its object, merely our internal consciousness, and which we obtain solely through the medium of our inner consciousness. In contrast with this, I will define self-knowledge as the set or part of the set of information which we have about ourselves. This information is not limited to information about our internal consciousness and we can acquire this information by consulting other sources than our inner consciousness alone.

According to a classical view on self-consciousness (which I will criticize) examples of self-consciousness as above defined are for example the sensation of pain, the realization that one is happy, and the memory of a thought one had earlier. In all these cases we appear to be conscious of something that belongs to our consciousness: respectively, the feeling of pain, the sentiment of happiness, and a thought. It also seems to be merely our own consciousness which gives us access to these instances of our consciousness.

In contrast to this, an instance of the self-knowledge, as above defined, could be our knowledge that we were naughty as a child. For this to be self-knowledge, we need not remember this. It belongs to the information which we have about ourselves, but this information could come from what others, whom we trust, have told us about ourselves.

The definitions which I give of self-consciousness, on the one hand, and self-knowledge, on the other, are particular. They are neither meant to replace other existing definitions, nor to exhaust the meaning of self-consciousness and self-knowledge. I need, and merely use them, to argue that often when we, in my terminology, define something as self-consciousness, this in fact concerns, again in my terminology, an instance of self-knowledge: often when we think that we gained some information about ourselves through our inner consciousness, we in fact received this information through other channels than just that of our inner consciousness.

III. Experiential memory: a consciousness of consciousness, or of more than that?

Experiential memory is one example of a phenomenon which is, contrary to what is often thought, not an instance of consciousness of mere inner consciousness. Below, I will first illustrate what it means to think of experiential memory as if it were a consciousness of mere inner consciousness. This will simultaneously demonstrate that we are often tempted to think that it is just this. Secondly, I will enumerate factors that influence us to think thus. Thirdly, I will give examples of memories that are experiential, but are not instances of a consciousness of mere inner consciousness. Subsequently, I will explain what the former illustrations teach us about experiential memory and, more specifically, in which ways experiential memory proves not to be a consciousness of mere inner consciousness. I will

conclude with a note on what it means that experiential memory is, in my terminology, an instance of self-knowledge, rather than of mere self-consciousness.

1. Experiential memory is often thought to be a consciousness of consciousness

Experiential memory is the memory we have of experiences we personally had. It is often distinguished from the semantic memory we may have of facts, as well as from habitual memory. My memory of having been on a week-long boat trip once, is an experiential memory. My memory of the fact that there are Mayan temples in Mexico is a semantic memory. It is a memory of a fact I did not witness, but once learned about. When I habitually remember something, I remember how to do something. For instance, when you jump on your bike and you immediately take off, you are able to do so, because you remember how to bike. Your memory of how to bike is a habitual memory. In this paragraph about memory, I will only be analyzing experiential memory.

Experiential memory is often considered to be an instance of a consciousness of mere consciousness. It is then perceived as a representation of a past experience from the same perspective as the remembered event was perceived in the first place. Those who have this idea of what an experiential memory is, do not necessarily hold that we, in experientially remembering something, fully relive the original experience. They can recognize that we typically know when

we *remember* a *past* event and are not so absorbed by its representation, or fooled by the perspective of this representation, that we experience it as if it happens now. It is merely assumed that when we experientially remember something, and are aware of the fact that we *remember* it, we still represent it from the same perspective as we experienced it before. When I experientially remember how I once attended a Thanksgiving dinner, I am supposed to see a picture with my mind's eye and see the room and dinner table from the same perspective as I did then; I am supposed to still smell the food and hear how the voices of my family members sounded.

Those who think of experiential memory in this way, do so because they assume that we, in experientially remembering something, merely become conscious of our inner consciousness: I am assumed to have the just mentioned perspective and sensations in my experiential memory of the dinner, not because the dinner is still taking place, but because I am conscious now of what I was conscious then. The fact that I personally experienced something, makes one expect that I still have personal access to that experience.

If this were the right way to conceive of experiential memory, then experiential memories would resemble retentions. Let's think of our retention of a note. We can hear melodies, because we still retain the sound of the previous note when we hear the next note. The previous note no longer resounds. All we are still aware of, is the consciousness we just had of it. We can equally retain distinct sensations of a multi-sensational event. Say you suddenly slip while playing tennis. You fall on the ground, feel a pain flash through your wrist, and you smell and taste the gravel. When you recover from the first shock, during which your consciousness seems to have been blocked for a moment, you can

experience how you retain all these sensations. You can feel how the second flash of pain in your wrist *followed* the first, how the smell of the gravel is *still* there, and how some of the gravel *stays* sticking to your lips. Here too, you are able to relate a subsequent experience to a previous one, because you are still conscious of the sensation (i.e. consciousness) which you had before. These retentions are similar to how many picture experiential memory. In both cases, the experienced event passed, but our consciousness is supposed to still give us access to the consciousness we had of it. According to those who think thus, the only difference is that (we are aware that) the retention takes place right after the experienced event, whereas (we are aware that) there is typically a bigger gap between the original experience of an event and our experiential memory of it.

Some philosophers, who expressed that they conceived of experiential memory in this way, are Russell, Judson and Alexander of Aphrodisias. Russell wrote that

> "[i]t is obvious that we often remember what we have seen or heard or had otherwise present to our senses, and that in such cases *we are still immediately aware of what we remember*, in spite of the fact that it appears as past and not as present" (Russell 2001 [1912], p. 26, my italics).

I will take issue with Russell's quick inference that when we remember something because we were acquainted by it, we are still *immediately* aware of it. Even when our own past experience allows us to remember something, we may no longer have an immediate awareness of this event. Our memory may present itself as a piece of propositional knowledge, rather than as a perceptual awareness of this event. This propositional knowledge does not have to be immediate. It

could be mediated by the story I started telling about the event, or by the new context in which I place it now. In such a scenario, I still only remember the event, because I was present at the event. Yet, the way that I remember it, is mediated by the way I started looking at it afterwards.

Judson (1987-1988) slightly adjusts what Russell says. She agrees that in episodic memory we are presented with a picture of what we remember, but argues that this picture is not necessarily what we saw in front of our eyes when we first experienced this event:

> "If I am asked, 'What colour was the sea that afternoon?'... I may find myself picturing the afternoon's sea...and forming judgments on this basis; or I may, alternatively, get an image which 'feels right' as an answer to the question...The judgment...requires the image...but it is not based on it" (Judson 1987-1988, p. 68).

I will admit that we can form a picture of what we experienced in experiential memory, and agree with Judson that this can be a newly generated picture, rather than the image we had access to when we first experienced the event. However, I will deny that an experiential memory requires an image. We can remember that we did something without visualizing it.

Still, what Russell and Judson say is seconded by many others. Just as Judson, Alexander of Aphrodisias conceives of our memories as images. He assumes that they are similar to the images we create in fantasy:

> "Now what fancy or imagination is, we may explain as follows: We may conceive to be formed within us, from the operation of our senses about sensible objects, some impression, as it were, or picture in our original sensorium, being a relic of that motion caused within us by the external object; a relict, which, when the external object is no longer present, remains, and is still preserved, being as it were

> its image, and which, by being thus preserved, becomes the cause of our having memory: Now such a sort of relict, and as it were impression, they call fancy or imagination" (Alexander of Aphrodisias 1765, Book 3 358 n3, p. 135).

Don Locke, in his turn, agrees that when we remember something personally, we can still feel how we experienced the remembered event at first:

> "personal memory consists in bringing some previously experienced thing to mind, thinking about it again, and going over *what it was like*" (Locke, 1971, p. 76, my italics).

2. Influences on our conception of experiential memory

Different factors can lead up to the thought that experiential memory provides us with something like a picture or movie of a previous event, shot from the same place as we were at during our original experience of this event. I will enumerate some of these. They illustrate how our way of thinking about experiential memory can be influenced by our definition of this memory (1-3), as well as by the way we speak and by our visual culture (4-5).

First, we judge that true experiential memories, in contrast to false ones, should sketch an accurate picture of what we experienced before. This can make us think that to truly remember a personal experience, we should remember it as we experienced it then. For what could be a more accurate depiction of this experience, than an exact repetition of the vision we once had?

Secondly, experiential memories are per definition *personal* memories. They are memories of something *we* experienced.[44] This can make us assume that to have an experiential memory, is to be aware of our own consciousness, as well as that in experientially remembering a previous experience, we review it from the same perspective as we saw it before. For when we think about how our personal experience of an event differs from the experience that others may have of it, we may conclude that others can have similar feelings about the event as we do, but that they can never feel our feeling about it. We each have our own consciousness of the event. Further, others

[44] It is true that Parfit (1984, p. 220-223) has challenged this and tried to convince us that memories could be seen as a sub-category of the larger category 'quasi-memory,' under which, both personal memories could be subsumed, and something like personal memories, which are not, in fact, personal. However, even Parfit admits that this is just a logical construction meant to show that if memories are seen as part of a larger category, they do not have to be seen as necessarily personal. So, we could talk about human characteristics without using the concept of a person. In a personal conversation during the academic year of 2009-2010, Parfit stressed that real memories, memories as we factually have them, and which we are willing to call memories, are personal. His extra category of a phenomenon, which would be exactly like a memory, but then not personal, is not meant to replace our current category of memory. Instead, it is a category next to that of our category of memories. We then have a category of memories which are personal, and a category of something like (but not) memories which are not personal. Both of them then belong to the overarching category of *quasi*-memories.

Memories are not necessarily personal. When you are unsure whether you remember something or have made it up, your criterion for it to count as a real memory is that you have really experienced what you now seem to remember.

For an interesting discussion on why it is not even possible to conceive of a broader category that includes both memories and something just similar, but not personal, memories, cfr. Wiggins (2001, pp. 193-225). Wiggins establishes why Parfit's attempt to define quasi-memories, in a no-identity-implying-way, fails. Parfit (1984, p. 220) defines a quasi-memory as a phenomenon, in which I seem to remember a past event, someone experienced this past event, and my memory is causally dependent, in the right way, on the past experience. Wiggins points to the fact that Parfit did not find a way to define this 'causally dependent in the right way,' in a way which does not imply identity. What we would normally regard as 'causally dependent in the right way' would be 'caused by our original experience of this event'.

cannot have exactly the same perspective of things as we do. If they witness the same event, they see it from an at least slightly different angle. Having a memory of a personal experience may then seem to mean having a memory of the unique features of this experience that make it personal, i.e. our own sentiments of this event at the event and our own perspective on this event at this event.

Related to this, it may just seem to be impossible to remember, in experiential memory, the remembered event from another perspective, than we originally experienced it – say from my friend's perspective, in which I too would be visible in the picture. I could possibly envision such an image, but we would call that a fantasy or hallucination, rather than an experiential memory.

Fourthly, our way of thinking of experiential memory may be influenced by (other than purely defining) ways, in which we talk about it. For instance, we say that 'to recall a previous experience' is 'to call this previous experience to mind'. When we take this image literally, we will think of experientially remembering as letting an experience reappear, just as we can make a friend return by calling him. We then think of a memory as a previous experience that appears again before our mind's *eye*.

Lastly, our visual culture can equally influence the way one thinks about experiential memory. Cartoonists mostly show that a character remembers something by drawing a thought-balloon above its head, with a depiction of the remembered event in it. Film-makers often depict memories as flash-backs: they film and show the remembered experience from the perspective from which the character is supposed to have seen it. For the artists, these are just two easy ways to depict memories, but seeing these artificial visualizations of memory

regularly can make us think that our own memories are exactly like this, i.e. that we, in experientially remembering something, see an exact copy of what we saw before.

3. Experiential memories, in another guise

While it is easy to see how this specific way of thinking about experiential memory may come about and be reinforced, it is also quite evident that it does not correspond to how our experiential memories generally present themselves to us. I will now give a couple of examples of experiential memories in another guise than the one just discussed. I will analyze them in the subsequent paragraph. The instances of experiential memory that I will discuss are memories of personally experienced events. Yet, these events are (1) neither relived in a way similar to how they were originally experienced, (2) nor are they reviewed from the same perspective as they were initially seen. The examples are real and personal. My hypothesis is that your own experiential memories often present themselves to you in a similar way.

(1) When I have an experiential memory, I do not typically relive it. When someone says 'remember how funny that was', I mostly remember that it was really funny and sometimes giggle about it again, without re-experiencing the fun I had then. Often, I even remember the event as funnier, or as less funny, than I originally judged it to be.

(2) In most of my experiential memories I just remember that I experienced something. I do not always visually display the

remembered event and, when I do, I seldom review it from the same perspective, from which I originally observed it. For instance: I remember that I once climbed a rope and that it snapped. I remember that I suddenly crashed to the ground, without having foreseen this and that I felt the impact go through my body. I also remember that I had a headache afterwards, and that my mother felt responsible and sorry for the faulty rope. I remember all of this from experience, not from stories that others told me afterwards. Still, I do not remember what the garden looked like from the perspective I had then: I do not remember whether I, from the rope, first saw some poles or the trapezium. In fact my memory of this event does not at all present itself as some kind of image that I see from the same perspective as I saw it then. Nor do I remember how I felt the pain go through my body. Did it start from my lower back and then shoot up? Or did I immediately feel it in my head? Did I feel it at all in my neck or teeth? I could not say. I remember that I had pain, but I do not remember any details about the pain and, in remembering the pain, I definitely do not re-experience it. I really just remember that all of this happened and some pictures flash through my mind of where this happened and with which kind of rope. But these pictures do not stem from that particular event. When I look at these pictures, I do not see them from the perspective I had just before the rope snapped. I just have them because the rope was there for years.

I also remember how twice I searched for my mother, in vain. I remember how, in one case, I climbed onto the slide in the garden, to look over the neighbours' wall, to see if my mom was there. The strange thing is that I, in remembering this, seem to picture myself. I seem to see what I wore: it was a white t-shirt with a Mickey Mouse

logo on it and I was not wearing any trousers. I see myself, standing on the slide and getting of it in a careful way because I was barefooted. Again, I do not just imagine this event after somebody told me about it. I remember it from experience. But, even though this is an experiential memory, I cannot be said to have the same perspective in remembering the event as I had it when I first experienced it: I now have some picture of me standing on the slide, whereas at the moment I was standing on the slide, I must have seen the neighbours' garden.

Of the other time I searched for my mother, I remember that my brother and I ended up crying in front of the big mirror of the hall. I also remember that my mother gave us a bonbon when she came home, and that we all sat in the living room for a while. Still, while I remember this event and the quite exact places of where I then found myself (first in front of the mirror and then on the couch), I do not review the realms I was in, as I saw them at the time. I do not remember seeing my mirror-image in the mirror, and I do not remember seeing the living room from where I sat. I remember that I was located at a particular place, but, in my memory, I do not review the world as I saw it from that place.

A last personal example: I remember, from experience, that as a child and teenager, I stood in front of the bathroom mirror a couple of times, and looked at myself intensively. I was trying to memorize the picture of myself. I wanted to find out whether it would be possible for me to later remember exactly how I looked at these different instances. I did not turn out to do so. Although I remember that I looked in the mirror and tried to memorize the mirror-image of myself, I no longer remember what this mirror-image looked like. I do not recall how my face must have stared back at me.

These examples of experiential memories are not meant to demonstrate that we never remember something from a similar perspective as we first experienced it. People do have flashbacks where this seems to be the case. We can also achieve to slowly and deliberately regain a past perspective of things[45]. My sole aim was to illustrate how an experiential memory does not always present itself as a re-experiencing or visualization of a past event, perceived from the same perspective, as it was perceived first. I expect that when you reflect on your own experiential memories, you will equally recognize that they often consist in the realization that you experienced something, rather than in a reliving of the experienced event or in reviewing it from the same exact perspective as you experienced it at first.

Philosophical voices that agree with what I have contended, are those of Bernecker (2007) and Martin & Deutscher (1966). Bernecker bases himself on scientific research:

> "… the frequency of memory images varies greatly from one person to another (Nigro and Neisser, 1983). Some people report that their mental lives are replete with imagery as vivid and detailed as the actual scenes they recall. For others, imagery is uniformly vague, dim, and fleeting. And, for a final group, there seems to be no imagery at all" (Bernecker 2007, p. 138). Also: "… many

[45] Sometimes you remember something, but have to retrace your steps to remember exactly what happened: something makes you think of an earlier event, but to bring this event fully to mind, you have to ask yourself some questions such as 'where did this happen again, when, and with whom'? You may then be able to bring this event back to mind. But what you recall is already more than the memory you just had. You arrive at it by reconstructing a certain event, but this exact reconstruction was not necessary to remember the event in the first place.

autobiographical data are remembered by description" (Bernecker 2007, p. 139).

Martin and Deutscher argue for something among these lines. They acknowledge that many report to visualize, or hear again, what they remember (Martin and Deutscher 1966, p. 164-5), but they equally recognize, just as I will do, that experiential memories can also take the structure of semantic memory. About this case, Martin and Deutscher (1966, pp. 162-3) say that we remember *that* we personally experienced something without, in many cases, remembering much of the details of this experience.

4. Experiential memory is more than consciousness of mere inner consciousness

We have now seen that often, when someone experientially remembers something, a correct description of this memory is not that he still feels how he felt something, or sees how he saw something, but, instead, that he still knows that he experienced this.[46] Translated into philosophical terms, one can say that the structure of experiential memory, in these cases, resembles that of semantic memory. In this experiential memory with semantic features, we think of the remembered fact as a third person could think of it: we do not re-live or review it as we originally did, but look at it from the outside, and

[46] Of course, you may just seem to know something, but, in fact, be mistaken. It also happens that we are not exactly sure whether we remember something that actually happened, or whether we are just remembering the image that we created for ourselves after we were repeatedly told about or heard about a certain event.

can even picture how we looked at the time, instead of just seeing how the world around us looked from the viewpoint we had then. This is so, all the while we personally experienced the event and we remember it because we personally experienced it. Picturing this event is not even necessary. Just as I can semantically remember it being a fact that Hitler had a dog and may, in remembering this, either picture him and his dog or not, I can experientially remember *that* I was upset when our family dog was put to sleep and picture the whole scenery (myself included), or just remember this as having been the case. In contrast to the former experience, I personally – and so experientially – experienced the latter. Yet, both memories can present themselves in the form which one often solely associates with semantic memories.

The semantic form which an experiential memory often takes, can be one of the factors that sometimes causes us to be confused about the origin of a particular memory. Sometimes we know of a past event in which we were involved, both because we personally experienced it, and because we heard about it afterwards. When we remember this event, we may fail to determine whether we experientially remember this event in a that-format, or rather semantically remember what someone told us about it.

Still, when we experientially remember something, we typically know that what we remember really happened. How do we know this? With what we have seen so far, we can rule out two explanations. Both explanations work with the idea that we experience our personal memories as real and personal memories because our consciousness gives us a clearly personal access to it. The first explanation says that we assume that an experiential memory is a memory of our past because it has an intensity, which is faint enough not to be perceived

as a current experience, yet vivid enough to make us suppose that we experienced what we seem to remember, rather than having imagined it, or rather than thinking back of something someone told us about. According to the second explanation, we experience our experiential memory as a memory of something that we personally witnessed because the perspective that, in our memory, we have of the remembered event, is the perspective that the subject, who experienced this event, must have had; it is believed that we can only have this perspective now, if we personally experienced what we remember. These explanations can be excluded because our experiential memories present themselves with all kinds of vivacity (they are sometimes more and sometimes less intense than the original experience), and because we do not always look at a remembered event from the perspective, from which it was originally experienced. My alternative to these explanations is that, in remembering a personal experience, we know that we really experienced the remembered event, because we knew that we were really experiencing it at the moment of the original experience, and, at that moment, stored this information about the reality of being involved in the event. In the remainder of this paragraph, I will analyze this dispositional knowledge which we recall in experiential memory, as well as the framework that is required to do so. This should demonstrate that experiential memories are not merely instances of mere inner consciousness.

When we return to the experiential memories, which I described in the previous paragraph, we can see that the following are some of the things that I remembered because I already knew that they were the case at the moment of the original experience: how I was a subject for

myself; how I was a subject for others; how I was a material object in space; and how my experience took place at one particular time. Take my remembrance of falling to the ground when the rope snapped: I remember that I was in pain, and so remember that I was a subject for myself; I remember that my mother was worried about me, and so remember that I was also a subject for others; I remember that I was in the back of the garden where I was visible to others and was subject to the rules of gravity, and so remember that I was a material object; and I remember that all of this was the case at one particular dry day, and so understand that it all happened at one particular point in time. This requires that I am aware of more than my inner consciousness alone – both at the time of my memory and at the time of my initial experience. (1) Firstly, I can only experience and remember this, if I do not just assume that my experiences take place in an inner time constituted by my own consciousness, but also in a time and space that I share with others. (2) Secondly, I can only experience and remember this, if I can imagine how what I experienced is seen from a second and third person's perspective, rather than just from a first person's perspective, which a mere consciousness of one's own consciousness provides.

(1) Let's first zoom in on how, in the above described experiences and memories, I locate myself in time and space. I, there, locate myself both in an inner and in an intersubjectively shared time and space. An inner time and space can be constituted by my consciousness of my own consciousness alone. Intersubjectively shared space and time is a space and time in which both, I, and others like me, could locate me. For instance: when I fell down the rope, I located myself in an inner time and space when I felt how I suddenly had a pain that was not

there before, and when I felt how this pain flashed, i.e. how it started somewhere and then rose or descended through my body. To experience all this, I need not know that there is a time, which we divide into seconds, and in which we can so make appointments with others. Nor need I know that there is a space in which material objects can occupy a place. All I need to be conscious of is my own consciousness. The detection of change in this consciousness is what gives me the idea that there can be a before or after, as well as that there are distinct places at which I can sense things. I located myself in an intersubjectively shared space and time when I realized that it would take my mom some time to reach me: in estimating the distance between me and my mom, I located us in the same space; in thinking that she would reach me somewhat later, I situated us in the same time. The same is true for the situation in which I now find myself, and in which I remember my previous experience: I develop thoughts and perceptions that make me both locate myself in an inner and an intersubjectively shared space and time.

The fact that I locate myself in this intersubjectively shared space influences my experience of experiential memories. When I experientially remember something, part of the characteristic of this memory is that I take the remembered event to have taken place at one particular moment in my life. I may not recall what this exact moment is, i.e. what the date was or whether it happened before or after another particular event. Still, I assume that this remembered event must have one particular place in the order of events that I experienced. I only assume this because I already have the idea that I live one particular life and only have one particular life-history, in which all experienced events must have their place. This very memory cannot, on its own,

give me this idea. Taken into isolation, the remembered event could have occurred at multiple times and even at multiple places at once. Nor can the totality of my consciousness of inner consciousness give me the idea that the remembered event must have had its particular place in my life-history. In my becoming conscious of consciousness, all kinds of conscious experiences come to mind. Some appear to be related, but whether there is any relationship between many others remains unclear: the content of consciousness does not inform us about a general order where all conscious experiences must have taken place, nor does the nature of consciousness determine that there must be such an order at all. We can have multiple conscious experiences at the same time and we can jump from a thought of being at one place, to a thought of being at a radically different place. So, as far as the nature of consciousness is concerned, there is no particular reason why one specific instance of consciousness should have its chronological place *between* two others. It is only because we already have the idea that we have just one particular history, in which we were always just at one place at a time, that we assume that a remembered event must have occurred at one particular moment in our history.

Having this idea requires that we situate ourselves in a time and space that is not just inner, but intersubjectively shared. In being aware of our inner consciousness, we only situate those experiences in a time and space that reach as far as our retentions of experiences reach. In experiential memory, however, we assume that we experienced something at a particular moment in time and space, even when we no longer retain an uninterrupted series of events that we experienced between the remembrance of the event and the moment at which we first experienced the remembered event. We assume that

the event occurred at a particular moment and place in the time and space that we share with others. This is the time and space, in which we can all situate the same object that moves from one place to another, over a specific time-stretch. When I can no longer use elements of my inner consciousness to define when something I experientially remember occurred, I can still imagine and decide to trust that it occurred at one moment, to which one can refer in this intersubjective time and space in which we can all locate the same events. This is necessary for my experiential memory to get its specific character of having happened at an earlier moment in a time of which I no longer have any retention. When we understand this, we must conclude that an experiential memory is more than a consciousness of inner consciousness. As it appears, we locate many of the experientially remembered events, in a past, of which we would have no idea, if we were locked up in an inner consciousness, but, of which we can have a conception, because we also share a time and space with others in that we can locate the same events that could, in principle, always be witnessed by someone.

The fact that we have a conception of an outer space where material objects and we, as psycho-physical realities, can have a place, does not only allow our experiential memory to have the specific character that it has in as far as it makes us locate the remembered event in an intersubjectively shared time. It also allows us to distinguish (a) real experiential memories from false ones, and (b) thoughts of current experiences from non-actual ones. Without these distinctions, we would not conceive of experiential memories as we actually do.

(a) If I did not have the idea that there is a space that is not just constituted in, and by, my inner consciousness, I would miss a criterion

to distinguish between an experiential memory of witnessing an event and a memory of a mere (day)dream of witnessing this event. Herewith, my ability to experience an experiential memory as an experiential memory, i.e. as something I really experienced, would disappear. This, again, shows that when I do experience something as an experiential memory, more than a consciousness of my mere inner conscious is involved.

(b) This ability to distinguish between an outer reality, in which I really find myself, and a reality that I only represent in consciousness, further helps me avoid getting completely absorbed by my memory. It facilitates distinguishing between something I just remember, or fantasize about, and something that I actually experience, now, (i.e. which also has its place in the time and space that we share with others). For instance: when I sit on the couch and look at the inclement weather outside, my realization that this is the real outer reality, in which I find myself, allows me to understand that the picture that I envision of green Italian hills is just a memory or fantasy of these hills. This contrasts what happens in my dreams. One of the reasons why I can think that I really experience what I dream is that I am, at that moment, no longer aware of the sleeping room where I really find myself.

(2) So far, for the importance for experiential memory of my background knowledge that I live in a space and time which I do not just constitute by being conscious of my inner consciousness. I also mentioned that, in having an experiential memory, I can think of myself from a second and third person's perspective, rather than just from a first person's perspective, which a mere consciousness of one's own consciousness provides. For instance: when I see myself standing

on the slide again, I see myself from a third person's perspective – I see myself, just as somebody who stood behind me could have seen me, or as a camera could have registered me. When I remember how sad and miserable my brother and I looked, and which feeling this evoked in my mother, I remember us from a second person's perspective: I remember how we looked to someone who cares about us and I seem to remember the feeling we triggered in her. I can only remember these experiences like this, when I am conscious of more than my inner consciousness alone. To have these remembrances, I should at the moment of my initial experience already have been aware of the fact that I am, amongst other things, a material object, at which one can look from different perspectives, and not only a subject for myself, but also for others who can feel for me. It is because I already had this awareness then, that I can now shift perspectives on the remembered event. So, for my experiential memories to present themselves as they do, I have to know, both at the moment of the initial experience, and at the moment of my memory of it, that there is a world outside my own consciousness, in which I have a place, and that there are subjects with a different consciousness than my own.

Now that I have shown that our experiential memory – which may, at first instance, seem to be nothing more than an occurrence of consciousness of one's inner consciousness – in fact involves the awareness that we are psycho-material entities that take up some space and can be seen and followed by others, it becomes, again, more comprehensible why we can identify a self through its body, even when we judge that our conscious experiences and self-consciousness are what is most essential to us. When the content of my experiences and

my thoughts of my experiences (e.g. my personal memories) are themselves partly constituted by my awareness of the fact that I, thanks to my body, take up one particular visible space, at one particular moment in time, then a conscious self appears to not just be a mental self. Many of its conscious experiences, and even of its self-conscious experiences, are informed by its existence as an intersubjectively observable psycho-physical being. As a consequence, fully describing a diachronic self involves describing it as a physical being which, as long as it exists, takes up some space in time. A self can only be the diachronic being it is, even in consciousness, if it is also a physically observable being in an intersubjectively shared space and time, and can think of itself as such.

5. Experiential memory as self-knowledge

The fact that we, in experientially remembering something, situate the event at a particular (although perhaps undefined) moment in our life history, does not imply that, so as to experientially remember something, you first have to check, every time, if what you seem to remember can have happened in your life, given the places you have been at and the route you have followed between them. Our experiential memory is permeated by the background assumption that, once in our life, we experienced the remembered event, but explicitly checking whether the latter was the case is not a necessary component

of the genesis of our experiential memory.[47] The assumptions that we must make to interpret memories as experiential memories, are just background assumptions. They determine the way we think about the world and ourselves. They are glasses through which we see things and, because of which, we can interpret memories as memories – but, they are not elements of which we first need to check the presence, to then determine that a certain thought is a memory.

It is true that one's statements about one's own past are grounded in the assumption that one is bodily continuous with the person one is talking about: one was always where this person was. In this sense, I disagree with Shoemaker who says that

> "[o]ne's statements about one's own past, when made on the basis of memory, are not grounded on bodily identity, or spatiotemporal continuity, as a criterion of personal identity; they are not grounded on the knowledge of *any* physical relationship between one's present body and a past one" (Shoemaker 1963, p. 34).

I argued that, when we, in remembering, state something about our past, we assume that we experienced the remembered event at one particular time and at one particular place. This is so, even when we do not recall the exact time and space of the remembered event. We assume that we, being where we are now, were then at another

[47] I say that it is not a *necessary* component in the genesis of this memory, because it can occasionally be a component of this memory. It, for instance, functions as such a component when a thought comes to mind, and you do not immediately know whether this is a thought of something you really experienced or a thought of a dream you just had. In this case, you can try to see if you remember other things, such as where you were the previous day, to see whether it is likely or implausible that you have really experienced this. It can then suddenly appear to you that you did, in fact, experience this. In this case, checking whether something really happened has a part in the genesis of an experiential memory of yours: it gives the thought of the supposedly experienced event the character of a memory, rather than that of a good or bad dream.

particular place. We only assume this, because we learned that we are continuous beings that take up some space in a material world and that we have always been at one particular place, at one particular time. The latter is part of our knowledge about the world and a prerequisite for it.

Still, I agree with Shoemaker that

> "[i]t would be absurd to suggest that in order to be entitled to say 'I remember taking a walk last night' I must first examine my body and satisfy myself that it is the same as some particular body that existed last night" (Shoemaker 1963, p. 34).

It is, indeed, not so that in order to be entitled to say that you remember something, you must first have checked whether you were really once physically present at the event that you seem to remember. Although this must have been the case for this memory to be a true memory, and although you assume in remembering that this was the case, you do not have to check, every time, that this was the case, before you can say that you remember something. Still, if we would coincidentally find out that we were not physically present at the event that we seem to remember, we would conclude that what seemed to be a memory was not a memory after all. This means, after all, that our reliance on memories of our past depends on a physical criterion for personal identity.

Shoemaker was convinced that this could not be the case. He thought he had the ultimate argument to prove that our reliance on memories precedes any testing on whether we were physically present at the event that we seem to remember, when he said that we can only test the latter by trying to find out, through memory, whether this was the case (Shoemaker 1963, p. 33, 35, 39). This is a fallacy. We could,

just as well, learn from others that we were not physically present at a particular event, or we could use our reason to conclude that we could not possibly have been at a certain place.

By now, I hope to have shown that our experiential memory, sometimes thought of as a consciousness of our own inner conscious only, is, in fact, also informed by our idea of the spatio-temporal world we live in and the place which we have in this world. In the next chapter, I will illustrate how this is equally so for our sensations, and even for the thoughts that we have about ourselves qua mental selves. Together, this should illustrate how different instances of our supposed consciousness of inner consciousness are, in fact, not a consciousness of inner consciousness alone, but are equally influenced by our assumption that we take up a specific place in the spatio-temporal world. This should help us understand why we can identify a person by his body, even if we judge that his psyche is what defines him most: even in experiencing himself as a mental being, a person equally assumes himself to be a physical being.

CHAPTER 6:
ASCRIBING SENSATIONS AND THOUGHTS TO ONESELF

What holds for experiential memories also holds for sensations and thoughts about myself qua psychological self: while, in having them, we may just seem to be conscious of our inner consciousness, more is involved.

I. SENSATIONS

At first sight, we seem to access sensations through consciousness – if we would not have consciousness, we would not become aware of any sensations – and sensations seem to be nothing, but an awareness of a current conscious experience. Yet, just as is the case for experiential memories, these first ideas we may have about sensations do not account for their full phenomenal truth. I will show how the way we ascribe sensations to ourselves, and even the way we experience them, is informed by our idea that we live one continuous life, through one body. I will do so, (1) by demonstrating that we can only (self-)ascribe a particular sensation if we can attribute different sensations to the same located self, and, (2) by showing that how we experience a sensation depends on the image of the body we have.

If I am successful in doing so, this will, again, help to explain why our re-identification of a particular body can make us recognize a self

as the same self, even when we consider this self's consciousness as that, which is most essential to it. For, if the consciousness it has of being conscious is informed by the idea it has of living through a particular body, then fully recognizing it as a self-conscious being involves identifying it as a being, who has an idea of living through a particular body. Identifying someone by means of the body which has an important part in constituting her self-consciousness, does not, then, seem to be completely opposite to identifying her as a self-conscious being. This explanation does not sketch how we start re-identifying persons by their bodies. It merely shows why it is not absurd or irrational that we do so. Genealogically speaking, it at most explains why once we re-identify persons thus, we do not feel forced to stop re-identifying them in this manner.

1. We can only (self-)ascribe a particular sensation if we can attribute different sensations to the same located self

A first requirement for one to ascribe a particular sensation, as a particular sensation, is that one has, or has had, more than one sensation. The reason for this is twofold. (1) A particular sensation can only be isolated as a *particular* sensation, in the sense of being numerically or qualitatively individuated, if one can distinguish it from other sensations. (2) Sensations can only be distinguished from one another by a self who experiences all of these sensations. For one to be such self, one must have a sense of being a self. And for one to develop

this sense of being a self, one must have more than one sensation. Let me elaborate on this a little further.

(1) Since both individuation and meaning come about through diversification, a particular sensation can only be recognized as a *particular* sensation, and, thus, one among other sensations, if there are other sensations. We can only numerically individuate a particular sensation if we experience that there are other instances of similar or distinct sensations from which it can be distinguished; and we can only recognize that a sensation has a particular quality, when we are able to experience that there are sensations with other qualities. Imagine a being who is born in a state which we would call 'pain', but who does not have any other sensation. This being could not even imagine that there is something as relief from pain, and therefore not form a wish to escape this pain. He will not recognize that pain is one numerical instance of a sensation that can arise and disappear. As a consequence, he will not be able to individuate the experience, in which he is indulged, as a particular experience, and so he will not be able to self-ascribe it: for him there is nothing (particular) that he can self-ascribe.

(2) Sensations can only be distinguished from one another by a self who experiences them all. A sense of self, and so a self, only comes into being when this self distinguishes itself as experiencing subject from what is experienced. This, in its turn, can only happen when this self has more than one sensation, as was already remarked by Hume when he answered the following question negatively.

> "Suppose the mind to be reduc'd even below the life of an oyster. Suppose it to have only one perception, as of thirst or hunger. Consider it in that situation. Do you conceive anything but merely that perception? Have you any notion of *self* or *substance*?" (Hume 1978, Appendix, p. 634)

The reasons why we do not conceive of anything more here are multiple. If there is only one sensation, the sensations lack which would allow a self to have any feelings about the present sensation, let alone to act on it. This means that there is no possibility for a self to develop a sense of character or agency. Further, if there would only be one sensation, and no other impressions, there would be no possibility for a subject to locate this sensation at a particular place: he would not know of any elements that could allow him to construct the idea of a space with different places. Hence, there would be no possibility for the ascription of feelings to someone who feels them at a particular place and must, thus, be at that particular place. The sensation would be everywhere and so there is no possibility for a subject to become a particular subject by being somewhere in particular. Lastly, the lack of multiple sensations would prevent anyone from ascribing the one sensation to the one who feels it, in contrast to the one who does not feel it. For, if there is only one sensation, one cannot even start to imagine that there is something, such as the ability to not feel it, and, so, neither that there could be someone who does not feel it (in contrast to someone who does feel it).

To have the required more than one sensation, to be able to ascribe a particular sensation, a particular sensation has to be located in space and time. It cannot be constantly everywhere. For, if it was constant and everywhere, then we would again not be able to individuate it as a *particular* sensation. So, in being aware of a particular sensation, we are also conscious of it being felt at a particular place and time, and not at another place, at another time.

As we already saw, our identification of place and time is never purely descriptive, but always, at least in part, indicative. As we also

saw, this requires that we have a consciousness of being at a particular place and time ourselves. It then becomes clear that to ascribe a particular sensation, we must have had different sensations, we must have been able to locate these sensations, and we must locate ourselves at a particular place and time.

2. How we experience a sensation depends on our image of the body we have

In the previous paragraph, I focused on the requirements for *ascribing* particular sensations. I concluded that one condition of the self-ascription of these particular sensations is that we can locate ourselves in space and time. Our self-ascription of particular sensations, thus, appeared to involve more than a consciousness of our mere inner consciousness. Now, I turn to our *experience* of sensations and show how our localization of these sensations is equally informed by more than our consciousness of consciousness alone. I will argue that sensations, on their own, do not make us localize them as we do, but that our localizing of these sensations is informed by a conception that we already have of the shape of our body. This is a fact, not a necessary requirement. It can be illustrated by thinking about a finding in neuroscience.

Neuroscientists showed us that certain experiences that are normally provoked by our bodily interaction with the world can also be initiated by an immediate tampering of the brain. One famous example is that of the trembling of a finger. In normal circumstances,

one of your fingers may start trembling after lifting excessive weight, or when you are tired or nervous. Neuroscientists have shown that a similar trembling of the finger can be provoked when they immediately tamper with your brain, by means of placing, on your head, a special hat that sends magnetic impulses to parts of your brain. Given that this is possible, we can also imagine neuroscientists tampering with our brain in, such a way, that it would provoke a pain in our knee, similar to a pain we may feel there after a fall. The question is, whether the neuroscientist would also be able to make me locate this pain at the same place as I would normally do, if I did not have a prior conception of having a body. My hypothesis is that he would not be able to do this by merely invoking a particular sensation, and that I would only imagine this pain to be in my knee, when I already have an idea of what it is to have a knee.[48] If I never before saw my body and did not learn that it is continuous, I would not be able to locate distinct flashes of pain in the same body part or, for that matter, at the same place. I would have no background knowledge that would allow me to

[48] Gareth Evans (1982, p. 251) says something similar when he says that, as a brain in a vat, I could only imagine sitting on a bank with my knees bent, if I were embodied before and my experience of my own body and my perception of those of others taught me what it meant to sit on a bank with my legs bent. He reasons that our brain as an organ, in itself insensitive, cannot give us any sensation and that, if we never had these sensations, we could not think of how they would feel like: we cannot construct them in our fantasies. I go one step further. I say that even when a sensation could be provoked in a brain-in-a-vat by tampering with it, we would not be able to locate it and experience it as we normally do, if we never had a real experience of seeing what it is to have a shape and solidity. We would have no incentive or means to construct the idea of being shaped thus.

(re-)identify a place. I could, of course, always feel a pain close to where I feel an itch, but even that would not help me to localize this itch and pain: they could be felt together at different places. If I do not have an idea of having one continuous body with a certain shape and solidity, I cannot localize sensations at one particular place in space.[49] Of course, I would be able to feel something when the neuroscientist tampered with my brain, in this way. My hypothesis, so far, is just that I would not locate it in my one knee, as I would normally do, if I do not receive more input than this particular provoked sensation alone. My second hypothesis, in this regard, is that my particular experience of the sensation will differ, depending on whether I can or cannot locate it in my knee. Something similar happens when you have an undefined tummy-ache and the doctor later tells you which organ exactly provokes this pain. The fact that you do not know where exactly to locate the pain in the first case, but you do know where to, in the second, will confine it in the latter case and give you a different experience of it. It is, thus, not necessary for a sensation to be felt, that you locate it on or in your body, but given that we do so, because we have an idea of the shape of our body and thus of how it occupies space, our locating of this sensation does influence our experience of it. This means that our experience of sensations is influenced by more than our

[49] It would not be relevant, in this context, to suggest that, as a brain-in-a-vat, we could still ascribe experiences to ourselves as a mental self, rather than as an embodied self. For, we set out to discover what is required of us to experience a sensation, which is, per definition, a located feeling. The question here is whether our ascription and experience of this located feeling only depend on our inner consciousness and inner constructed sense of space, or whether it depends on more than that.

inner consciousness alone: it also depends on the idea that we have of how our body looks like outside of an inner space. So, contrary to what at first may have seemed to be the case, both our ascriptions and experiences of sensations involve more than consciousness alone.

II. THOUGHTS ABOUT MYSELF AS A MENTAL SELF

What is true for sensations, is also true for thoughts we have about our psychological character, such as the thought 'I am stupid.' At first sight, these thoughts seem to fully consist in a consciousness of inner consciousness, arrived at through consciousness. Still, they only present themselves as they do, because we are aware of more than our inner consciousness alone and, more specifically, because we are aware of the fact that we are located in a space that we share with others.

Before I demonstrate that this is the case, let me recognize why this, at first sight, may not seem to be so. When we think about ourselves as mental selves, we may, at first sight, not seem to think of ourselves as located beings, simply because these thoughts do not seem to concern something physical that is necessarily located. Sensations are different: feeling them on, or in, our body gives us information about our body. But the thought, 'I am quite a character', says nothing about my brain that produced it.

Still, I-thoughts about one's own mental character involve more than a consciousness of consciousness. Even in thinking of ourselves as mental beings, we think that we are somewhere. We attribute a place to this mental self in the physical space that we share with others. In his well-known essay 'Where am I?', Daniel Dennett (1998) argues

that this is so because we consider ourselves to be functional selves, and therefore locate ourselves where we receive and process input, as well as where we generate output. I will first highlight how Dennett illustrates this and then further elaborate upon his statement.

In 'Where am I?' Dennett (1998, pp. 310-323) imagines a situation, in which scientists put his brain in a vat in a specific lab and send his body elsewhere. In a first scenario, the scientists connect Dennett's brain with his body, by means of radio-links. In this way, his brain can still send commands to his body, and his body can still send received sensory information to his brain, just as they would do when nerves connected them to one another. In a second scenario, these links are severed and the former body of Dennett is left without sensations or capability of performing actions. In the lab, Dennett's brain gets connected to a new body. In a third scenario, Dennett learns that scientists have developed a software program which makes a computer run exactly as Dennett's mind would: if it receives the same input, it generates the same output. The computer that runs this program is a copy of the functional organism which Dennett's brain is. The scientists have experimented with connecting Dennett's former body to this computer, instead of to his brain. Dennett remained unaware of this while performing actions. He could not tell whether his actions were initiated by his brain, or by this computer program.

Dennett wonders where he would consider himself to be in these respective scenarios. Would he consider himself to be in his brain, in the rest of his body, in both, or in neither? Dennett's answer to this question shifts, depending on the scenario he imagines himself to be in, and the angle from which he looks at it. This can leave one with the impression that he does not give one final answer to the question of

his essay's title, and that he just does what the title of his essay bundle says: brainstorming. This would be the wrong conclusion. If I interpret Dennett's essay correctly, then he holds that we conceive of ourselves as functional selves, who are able to both receive and process input, as well as to generate output; he also holds that we situate ourselves at the place of this functional organism. This would explain why Dennett takes himself to be simultaneously in the brain and the rest of the body, in the first scenario: the brain is where he receives and processes input and the rest of the body is where he generates output. My interpretation of Dennett's answer to his question also explains why Dennett believes that he would continue to live through his own brain and new body, in the second scenario (where his brain gets connected to a new body), and why he would then still be the same person: he still functions in the same way. He has received the same input, continues to receive further input, processes it as he processed it before, and generates similar output, for similar input. Lastly, my interpretation of Dennett's answer to his question 'Where am I?' fits what he says about the third scenario. It explains why he thinks that he would still be the same self when the computer program takes over the task that his previous brain fulfilled: considered as a functional organism, the computer-body-whole is identical to the brain-body-whole.

Dennett has a point: we are functional selves and perceive ourselves as such. We see mental selves in, and conceive of mental selves as, beings who can receive information, process it, and generate an output in response to it. In doing so, we physically locate these selves. We see them where they listen to us, smell, speak, or act. For instance, we

understand the will of a mental self to work longer, when we see how she programs her alarm clock earlier for the next morning.

However, Dennett is too quick in inferring that the fact that mental selves are functional organisms implies that a mental self can stay the same self, even if her 'hardware', which she needs to receive input, process it, and generate output, is replaced by distinct non-continuous matter. I expect that most of us will equally be reluctant to accept this inference. To have been a particular self, one must have experienced what this self experienced. Dennett assumes that we can say that the computer experienced the same as Dennett did, and is hence the same self, because scientists provide it at a later moment with all the input Dennett once received. I contest that the computer here receives the same input as Dennett did. Dennett will have some memories of past events, whereas the computer will merely have information about these events.

I have used Dennett's essay, 'Where am I?', as a starting point, to illustrate how the functional character of mental selves makes us locate them physically. I have also argued that this does not imply that a mental self can be downloaded and uploaded in any distinct kind of material. I will now further elaborate on *where* the functional character of a mental self makes us locate it.

(1) A first place is our body, through which we often produce output.

(a) The body is a place where we give expression to our mentality. While others may not typically perceive our brain and while they think of our mental character, rather than of some neuro-physical processes, when they think of us, they still get much of who we mentally are by

looking at what we physically express. They get this from our facial expressions and bodily behaviour. They also address us as mental beings via our body: they look us in the eye when they want to address us, and they caress our body when they want to comfort us. From this, we ought not to conclude that our mental character is something hidden inside us, and that others can only guess, on the basis of our bodily behaviour, what our mental character is really like. Part of our mental character is just how we visibly behave, and part of our psychological interaction with others involves communication through touch.

(b) The body is also that with which we act on intentions. This makes Dennett more tempted to locate himself, in certain situations, in the remaining part of his body, rather than in his brain, which is nevertheless necessary for the production of his intentions: "How could I be in the vat and not about to go anywhere, when I was so obviously outside the vat looking in and beginning to make guilty plans to return to my room for a substantial lunch?" (Dennett 1998, p. 313) Dennett's example goes to show that, as mental selves who are functional, we are, among other things, agents, and find ourselves at the place of execution of our actions. There, we can establish full agency.

(2) Secondly, as functional, mental selves, we also locate ourselves there, where we process our thoughts. It is true, that we do not have any proprioceptive feelings of our brain, nor typically see how our brain processes our thoughts. Still, we assume that we, as mental selves, cannot be but where our respective particular brains are. This becomes clear, in thought-experiments about brain-transfers, when we judge that we will be where our former brain, rather than where our

former body, is. We have learned from experience and testimony, and so have, as background knowledge, that one stops thinking, when the brain one had lacks oxygen; that one starts thinking again, when there is a quick renewed blood flow through this brain; and that no-one suddenly feels that he is where another brain than he was using before is.

Let me demonstrate the role that this kind of background knowledge can have for the localization of ourselves as mental selves, by means of three illustrations.

(a) In 'Where am I?', Dennett (1998) writes about how he would experience himself to be catapulted back to a vat in a lab, when the radio-links between his brain and body would be severed and he would no longer see, hear, smell, or feel with this body. He would not have this expectation if proprioceptive feelings, alone, could make him locate himself at a particular place: the brain in the vat, where he localizes himself in this situation, still does not have any proprioceptive sensation. Dennett only forms the idea that he will be in the vat in the lab again, because he previously learned where his brain was located (scientists showed it to him), and because he previously learned that this brain is what gives him consciousness (scientists demonstrated this by switching the connection between his brain and body on and off).

(b) It occurs that the words, in which we capture a thought about a mental aspect of ours, do not immediately make it apparent that we localize ourselves in having this thought, while the specific meaning, which this sentence has for us, shows that we do. It can then be our background knowledge of our localization that makes us localize ourselves when we have this thought. For example, I can have the

thought 'my English is poor.' Someone could understand the meaning of these words, without thinking that I localize myself when I have this thought. Still, I might just be doing that. I could have this thought and mean that my English is worse than the English I heard, or even mastered, when I lived in the United States. This would mean that when I think 'my English is really poor,' I am also aware that I am, now, at a specific place, rather than at another. It is my background knowledge of having a particular life-history during which I was at particular places, and of knowing that I can only be at one place at one moment, that makes that my thought 'my English is really poor' can mean 'the English produced by the woman sitting here, trying to type out her dissertation in the library, is very poor;' 'I write quicker and speak more fluent than I did before I lived in the U.S.A., but I'm definitely not so skilled anymore as I was while I was in the U.S.A;' as well as 'what a pity that time and distance prevent me from being in the U.S.A. right now.' I then do not think 'what a pity that I cannot mentally be in the U.S.A. right now'. I think 'what a pity that the U.S.A. is too far for me to (physically) be there right now'. So, a thought that, at first sight, may seem to be just about myself as a mental being can, in fact, be tied to, and informed by, thoughts about the physical being that I am as well.

(c) I may daydream about staying in bed and having the opportunity to sleep in, rather than having to sit upright at the office. It can seem as if, in my conscious activity of day dreaming, I consider myself only as a mental being, i.e. as a fantasizing mental being. This is not entirely true: here, I only consider myself to be a fantasizing mental being, because I simultaneously have the background knowledge that I am also a physical being, and cannot really be in the

bedroom, while I am in the office. I also only have the idea that the bedroom really exists and is a couple of miles away from me now, because I know that I could physically go there, rather than just mentally imagine myself there. Here, my background knowledge about being a physical self allows me to understand in which kind of mental thinking activity I am involved, i.e. daydreaming about something that really exists, rather than a current experience of something, or a wild fantasy about a non-existing world. Again, a thought that may have seemed to be a thought about myself, as a purely mental being, appears to simultaneously be a thought about myself, as a physical being.

All these illustrations should have established that our notion of ourselves as a functional self, together with our background knowledge of how we function in this world, makes us locate ourselves as mental selves, even if mentality, at first sight, seems to be radically distinct from anything material and localizable.

III. SELF-KNOWLEDGE AND LOCATING ONESELF

Now, I have argued that, often, when we are first inclined to think that a specific instance of self-consciousness consists in a consciousness of mere inner consciousness, this consciousness is, in fact, equally informed by our knowledge of our diachronic physical existence. (a) We saw that a mere sensation does not give us a sense of self, and that we already have to have an idea of being a diachronically existing self with an integrated body, before we can individuate a sensation as a particular sensation. (b) We also saw that many

thoughts, that at first seem to concern our mental self only, in fact, also say something about us as diachronically existing physical entities.

However, it is not because we assume our diachronic and physical existence in these cases, that we are also always consciously and actively aware of this existence. It is as Evans says about someone's awareness of being in front of a house:

> "Certainly what he perceives comprises no element corresponding to 'I' in the judgement 'I am in front of a house': he is simply aware of a house. But if we are to interpret a judgement made upon this basis as having the content 'I am in front of a house', we must have reason to suppose that the subject regards himself as recognizing the existence of a state of affairs of precisely the same kind as obtains when, for instance, a car is in front of a house. So what he envisages, or judges, certainly comprises two elements spatially related, although what he sees does not. (This only goes to show that it is not a good idea, in attempting to determine the content of a person's judgement, to examine nothing but the content of the perceptions which can legitimately give rise to it)" (Evans 1982, pp. 232-233).

As I mentioned above, one should not conceive of the background knowledge here in play as a true belief which we explicitly justified. It is not the case that we first experience the world and subsequently infer that we must be diachronically existing physical beings. To state things thus, would give the impression that we already individuate ourselves as mental selves before we perceive ourselves as being housed in a physical body. I have shown in multiple ways how our initial consciousness of being one mental self is already informed by our knowledge of being one physical creature. Our knowledge, that we are diachronically existing physical creatures, is experiential knowledge: it stems from our experience in and of the world, but is not

arrived at by inferring something from this experience. Our assumption that we are diachronically existing physical beings always already shapes our experience of the world we live in.

IV. IMMUNITY TO ERROR THROUGH MISIDENTIFICATION

There is yet another occasion at which my self-consciousness may, at first sight, seem to be a mere consciousness of inner consciousness, but, upon further examination, involves more than that. This is so, when my reference to myself is immune to an error through misidentification. This reference to myself is such, that there is no way I could have been mistaken and referred to someone else when I said or thought 'I'.

There are two main reasons to think that our consciousness, which make us refer in this way, is merely a consciousness of our inner consciousness.

A first reason comes to light when one looks at Sydney Shoemaker's interpretation of these self-references which are immune to error through misidentification. Shoemaker holds that these self-references occur when my affirmation of the instantiation of a property is immediately an affirmation of the fact that *I* have this property (Shoemaker 1968, p. 565). This is, for instance, the case when I feel a pain: when I feel a pain, I am immediately aware that *I* am in pain; I do not have to infer this.[50] The opposite is true for the instantiation of the

[50] "There is no difference between believing that one is in pain and being in pain, so there can be no question of explaining how it 'happens' that one believes that one

property of bleeding. When two people are fighting, one of them could think 'I am bleeding,' because he sees, in a glass window, that he has blood on his forehead. This person could be mistaken. The blood on his forehead could be the blood of the person with whom he is fighting. Here, the person's affirmation of the instantiation 'is bleeding' is not immediately identical to his affirmation 'I am bleeding.' The person infers that the latter is the case, because he saw that the former was the case and because the context of the fight makes it likely, yet not certain, that this blood would be his. Philosophers have assumed that self-references that are immune to an error through misidentification must be references in which we ascribe psychological properties to ourselves, because it is easier to find examples of an awareness of a mental property, which is at the same time an awareness of the fact that I have this property, than it is to find an illustration of cases in which the awareness of a physical property is also an awareness that I have this property. The thought that our psychological properties are properties of which we have an immediate awareness through our inner consciousness, consequently made philosophers think that a self-reference that is immune to error through misidentification must be a reference that we make on the basis of a specific inner consciousness of a property.

---

is in pain only when in fact one is in pain" (Shoemaker 1963, p. 219). "…when I say 'I am in pain,' what I am acquainted with, or directly aware of, is just the pain itself, or just the fact of my being in pain. It is this direct awareness that justifies me in saying 'I am in pain.' If by 'criterion for saying that I am in pain' is meant something from which it can be inferred that I am in pain, then of course I use no criterion; but if by this is meant something that shows that I am in pain, then I do have a criterion; the pain, or the fact of my being in pain, is itself the criterion…" (Ibid).

Secondly, it is held that when we refer to ourselves in a way immune to error through misidentification, we refer to a formal self. Referring to a formal self is referring to a self, without thereby envisioning which character traits this self has. For instance, I refer to myself in such a way, that when I say in a conversation 'I think so and so,' I have correctly referred to myself (and not incidentally to my neighbour), without first having thought about which character I have, how I look like, or where I am. The fact that no image of myself comes to mind in this self-reference, has made philosophers assume that rules governing the use of 'I,'[51] here make us refer to the subject of speech (who could have any character), rather than to a specific person, in a specific body. We have lesser of an image of subjects of speech than we have of specific personalities or characters: a spoken word can be specific, but any subject could pronounce it; whereas a personality, or character, is recognized by specific features. The self of this self-reference is called 'formal,' because the reference is formal (it is rule-based), and because this self is in a way empty (we have no specific image of it). One still conceives of this self as a mental self because it is supposed to be the subject of speech. Philosophers then hold that our inner consciousness alone can make us refer to this subject, because our inner consciousness alone can make us aware of the fact that we speak; we do not have to see in a mirror that we speak, to know that we speak.

However, in spite of these first ideas that we may have of self-reference that is immune to error through misidentification, we, in fact,

[51] "One can choose whether or not to use the word 'I', but the rules governing the use of this word determine once and for all what its reference is to be on any given occasion of its use, namely, that its reference is to the speaker, and leave no latitude to the speaker's intentions in the determination of its reference" (Shoemaker 1968, p. 559). See also Campbell (1994, pp. 101-102).

often refer to more in this reference than to what we can become aware of through solely consulting our inner consciousness. In the following two paragraphs, I will first show how there are also self-references that are immune to error through misidentification and refer to ourselves as an embodied subject. Subsequently, I will show that, even in cases where the self-reference that is immune to error through misidentification seems to refer to a mere mental subject, we can only refer to this subject, if we also assume that we have a diachronic and physical existence in a physical world.

1. An IEM self-reference to an embodied self

When Wittgenstein and Shoemaker first discovered the phenomenon of self-reference that is immune to error through misidentification, they described it as a phenomenon that merely (in

Wittgenstein's case[52]) or primarily (in Shoemaker's case[53]) concerned the self-ascription of psychological properties. At first sight, this seems

---

[52] "In the Blue Book Wittgenstein distinguished 'two different uses of the word 'I' (or 'my'),' which he calls 'the use as object' and 'the use as subject.' As examples of the first of these he gives such sentences as 'My arm is broken' and 'I have grown six inches.' As examples of the second he gives 'I see so and so,' 'I try to lift my arm,' 'I think it will rain,' and 'I have toothache.' He goes on to say: 'One can point to the differences between these two categories by saying: The cases of the first category involved the recognition of a particular person, and there is in these cases the possibility of an error, or as I should rather put it: the possibility of an error has been provided for.... On the other hand, there is no question of recognizing a person when I say I have tooth-ache. To ask 'are you sure it is you who have pains?' would be nonsensical'" (Shoemaker 1968, p. 556; quotes within quote from Wittgenstein 1958, pp. 66-67).

[53] Shoemaker is slightly more nuanced than Wittgenstein. From the fact that some self-references are immune to error through a misidentification, Shoemaker does not immediately conclude that these must be references to a self as subject, which is, at the same time, a self to whom psychological properties are ascribed. He is aware of the fact that there have been some problems with the idea that a purely mental subject would be the subject of thought, and that the postulation of such a self does, therefore, not immediately explain to whom a self-reference, which is immune to error through misidentification, refers: "The most commonly drawn conclusion, of course, is that one's self, what one 'calls 'I',' cannot be any of the physical or material things one finds in the world. But as is well known to readers of Hume and Kant, among others, it is also widely denied that any immaterial object of experience could be the subject of thought and experience. These views lead naturally to the conclusion that 'I' does not refer, that there is no self, or that the self is somehow not 'in the world'" (Shoemaker 1968, p. 560). Shoemaker further considers the possibility that "my 'self' [as] a flesh-and-blood person", "could be accessible to me (itself) in a way in which it is not accessible to others, so that in knowing that what is presented to me is presented in this special way – from the

inside, as it were – I would know that it can be nothing other than myself" (Shoemaker 1968, p. 562). So, Shoemaker considers that a reference to oneself, as a bodily self, is immune to an error through misidentification.

Still, Shoemaker concludes that the primary self-reference, which is immune to an error through misidentification, is a reference to a self, considered as mental self. This is also the kind of self-reference he gives the most examples of – his preferred example being that of the self-reference in memory: "But the appropriate way of expressing the retained (memory) knowledge that at the time of its acquisition was expressed by the sentence 'I see a canary' is to utter the past-tense version of that sentence, namely, 'I saw a canary.' This, if said on the basis of memory, does not involve an identification and is not subject to error through misidentification" (Shoemaker 1968, p. 560). For an elaboration on the same theme, see also (Shoemaker, 1963, ch.4) and (Shoemaker 1970). Shoemaker reasons that this reference to a self, considered as mental self is the most fundamental form of self-reference that is immune to an error through misidentification, because other self-references could not take place if this kind of self-reference was not possible. "There, I think, is an important sense in which the 'use as subject' of the first-person pronoun is more fundamental than their 'use as object'" (Shoemaker 1968, p. 566). "Now there are M-predicates, e.g., 'is facing a table,' which can in some circumstances be self-ascribed without identification. But in order to describe the circumstances in which such self-ascriptions could occur and in order to formulate the grounds of such self-ascriptions, it would be necessary to employ predicates, P*-predicates ... The speaker of this language would have to learn to self-ascribe such M-predicates as 'is facing a table' under just those circumstances in which he would be entitled to self-ascribe certain P*-predicates, e.g., 'sees a table in the center of one's field of vision'" (Shoemaker 1968, p. 566). Also: "[I]f asked what it means to call a body 'my body' I could say something like this: 'My body is the body from whose eyes I see, the body whose mouth emits sounds when I speak, the body whose arm goes up when I raise my arm, the body that has something pressing against it when I feel pressure, and so on.' All the uses of 'I' that occur in this explanation of the meaning of the phrase 'my body', which in turn can be used to explicate the use 'as object' of the first-person

like a fair judgment. When I think 'I am happy', I may decide afterwards that I was, in fact, not really happy, but just content. However, it is implausible that I, in this reference, did not really refer to myself, but to another person. With 'I,' I refer to the subject of my thought and experience. More errors seem to be possible in ascribing a physical property to oneself. I first discovered that one can misidentify oneself in such a situation, when I once looked in the rear-view mirror of the car I was in, and thought that I was the girl with the short hair, whose image was reflected in the mirror. Shortly after, I realized that the reflection of the girl in the mirror was not a reflection of me, but of a friend, who was sitting next to me in the packed car. When I refer to myself as a material object, the chance of misidentification seems to be always there, just as it is always there when I identify other material objects. In this second case of self-identification, the possibility for misidentification was there, because I had to derive the conclusion that I am identical to the girl in the mirror from certain elements. I knew I had short hair for the first time in a long time, and that this made my face look slightly different than before. This is why, when I saw the girl with the short hair, I thought I saw myself. I identified the girl as myself on the grounds of several assumptions: namely, that I had short hair and looked slightly different than I looked before. In contrast, there was no room for me to misidentify myself, in the first case. I did not infer the thought 'I am happy' from some other assumptions and present factors. The thought and feeling, itself, immediately made me refer to myself.

---

pronouns in the self-ascription of M-predicates, are themselves uses 'as subject'"(Shoemaker 1968, p. 567).

In spite of this obvious difference between two kinds of self-ascriptions which could explain why the first is always, and the second never completely immune to error through misidentification, Gareth Evans (1982) and Quassim Cassam (1997) argued that a self-ascription of physical properties can equally be immune to error through misidentification.

Evans (1982, pp. 218-220) explains that we often do not realize this because we think of immunity to error through misidentification as something that applies to propositions. If this were so, then statements about our own body would, indeed, not be immune to error through misidentification. When I say 'I have goose bumps,' this proposition is not immune to error through misidentification. I could be looking at a picture of three girls sitting with their back to the photographer, seem to recognize that I am one of the girls, and derive from this identification that I had goose bumps there. I might be right about this, but it is also possible that the girl in the picture was not me, but just happens to look like me, sitting at a place where I once sat, in a similar bikini. In such a case, I would have mistakenly misidentified myself as the girl with goose bumps. Since this is possible, the proposition about my own body 'I have goose bumps' is not immune to error through misidentification.

Still, Evans would argue that the statement 'I have goose bumps' could be immune to error through misidentification. It is IEM when the utterance 'I have goose bumps' is not just a proposition, but also a judgment that came about through a specific channel we use to gain knowledge. Evans holds that when someone says 'I have goose bumps' and does so because he has the sensation of having goose bumps, this

judgment would be immune to error through misidentification.[54] He argues that this would be so, because, for me, there is no difference between feeling (and so knowing) that goose bumps are instantiated, and feeling (and so knowing) that I have goose bumps. This is just like there is no difference between the feeling (and so knowledge) that the property of pain is instantiated and the feeling (and so knowledge) that I am in a state of pain. According to Evans, then, immunity to error through misidentification applies to utterances about our own body, when these utterances are judgments and when we use specific ways of gaining knowledge to arrive at these judgments. So, utterances about our physical condition are not always immune to error through misidentification, but this does not exclude that they are, sometimes (namely, when we arrive at them through specific ways of gaining knowledge).[55]

---

[54] Evans gives a slightly different but, comparable example: "There is a way of knowing that the property of ξ's hair being blown by the wind is currently instantiated, such that when the first component expresses knowledge gained in this way, the utterance 'The wind is blowing someone's hair, but is it my hair that the wind is blowing?' will not make sense" (Evans 1982, p. 218).

[55] "Wittgenstein's discussion does not take sufficient account of the fact that the property of being immune to error through misidentification is not one which applies to propositions simpliciter, but one which applies to judgements made upon this or that basis. Once we appreciate this relativity to a basis, which arguably must be taken into account in the case of mental self-ascription as well, the fact that there are cases involving the self-ascription of physical predicates in which 'the possibility of error has been provided for' will be seen not to impugn the fact that there are cases in which it just as clearly has not" (Evans 1982, pp. 218-219).

Evans discusses "two ways we have of gaining knowledge of our physical states and properties, both of which give rise to the phenomenon of immunity to error through misidentification" (Evans 1982, p. 220). We just dealt with one instance of the first way. Our ability to feel our goose bumps belongs to what Evans describes as our "general capacity to perceive our own bodies" (Ibid). This capacity includes "our proprioceptive sense, our sense of balance, of heat and cold, and of pressure" (Ibid). As I just explained, Evans argues that it gives us knowledge because

> "[t]here just does not appear to be a gap between the subject's having information (or appearing to have information), in the appropriate way, that the property of being F is instantiated, and his having information (or appearing to have information) that he is f; for him to have, or to appear to have, the information that the property is instantiated just for it to appear to him that he is F" (Evans 1982, p. 221).

Secondly, there is the "way in which we are able to know our position, orientation, and relation to other objects in the world upon the basis of our perceptions of the world. Included here are such things as: knowing that one is in one's bedroom by perceiving and recognizing the room and its contents; knowing that one is moving in a train by seeing the world slide by (…) Once again, none of the following utterances appears to make sense when the first component expresses knowledge gained in this way: 'Someone is in my bedroom, but is it I?'; 'Someone is moving, but is it I?'" (Evans 1982, p. 222)

Quassim Cassam seconds this. He argues that, while a self-reference that is immune to error through misidentification is always a self-reference of a subject to itself as subject, this subject can still be a

bodily self. Cassam gives the same examples as Evans to demonstrate when this is the case:

> "This is so when a subject is aware of his sensations. One example of a sensation is the feeling of being solid. A subject for instance experiences its own solidity when it experiences the solidity of another object: Solidity can be either seen or felt, and an extremely important aspect of the perception of solidity by touch is the way in which it is bound up with a sense of the solidity of the perceiver. Solidity is typically felt as an impediment to one's movements, and to experience a solid object as an 'obstructive something' (O'Shaughnessy 1989, p. 41) is at the same time to be sensibly or intuitively aware of that which is obstructed – the subject of tactile perception – as something solid" (Cassam 1997, p. 52).

So, when a subject perceives that his body is solid, he does not just perceive that his body qua object (which he inhabits as a pilot) is solid. Instead, he feels how he qua subject is solid. In being aware of solidity, he feels the solidity of that, which is aware of solidity. This awareness of oneself as solid and, thus, as a body is immune to error through misidentification, for the same reason as the awareness of other sensations is; "if you feel a sensation, it is yours" (Cassam 1997, p. 64). This is also why "[t]hought-experiments according to which I would feel someone else's limb and thus wrongly say that I feel pain in this limb, don't prove that my self-ascriptions of pain are not immune to error through misidentification" (Ibid). You may be mistaken when you think you have a limb, but you cannot be mistaken about the fact that *you* feel that you have the sensation of having a pain in a limb.

With regard to the self-awareness that sensations give us, Cassam also points to something we saw before. When we have a sensation, we do not just locate it at the exact spot where we feel it. We are also aware that this spot has its place on a larger body. "[I]ndividual limbs

or body-parts are experienced as parts of an integrated totality" (Cassam 1997, p. 64). For "a limb with which one is not materially united cannot be a part of one" (Ibid, p. 65). This means that when we have a sensation "it is the bodily self rather than individual body-parts which is the presented subject of perception" (Ibid, p. 66).

A subject also refers to herself as a bodily self, in a way that is immune to error through misidentification, when she locates things in the her surrounding world. To locate where a specific object is, she has to know where she is, and so, where the object is, in reference to her. She cannot just locate objects by determining their place in reference to the place of other objects. In Cassam's words:

> "in egocentric spatial perception the objects of perception are experienced as standing in … spatial relations *to the perceiver* … [I]n experiencing objects as spatially related to one, one literally experiences the bodily self as located in the perceived world" (Cassam 1997, p. 53).

To be located, we must have some shape and solidity, by which we take up space. This is why, in locating other things and so also ourselves, we conceive of ourselves as bodily selves. If you think that you, as a bodily self, are where your point of view is, and refer to yourself in accordance with this, this reference is IEM. This is so because you do not infer, from certain perceptions, where you are, but your perceptions, instead, immediately instantiate you as a perceiving, perspective-having, and so located, subject.

In contrast to Shoemaker, Evans holds that judgments that are immune to error through misidentification are still expressions of knowledge. Shoemaker would deny that this is the case, since knowledge, for him, is a justified, true belief, and statements that are

IEM are not based on any criteria and, thus, not justified. Here, he is inspired by Wittgenstein, who defended that in order to identify something, the possibility of misidentifying this thing (i.e. mistakenly thinking that some identifying criteria are in place) should have existed. According to Shoemaker, there is no knowledge if there is no inference, just as, according to Wittgenstein, there is no identification if there is no process to be followed that can also lead to misidentification. Evans, on the other hand, explicates that our I-thoughts do represent a form of knowledge: they are expressions of some of our practical knowledge. Just as thoughts about the here and now, thoughts about ourselves give expression to "an element involving sensitivity of thoughts to certain information, and an element involving the way in which thoughts are manifested in an action" (Evans 1982, p. 207). For example, if I sense that I have a pain, I have access to the information that I have a body, of which one part now hurts, and this will typically result in me trying to remedy this pain.

This knowledge can be dispositional.[56] A certain pain may be so overwhelming, that we only seem to feel the pain without

[56] "Just as it is not necessary, if a subject is to be thinking about a place as 'here', that he actually have any information deriving from it, so it seems not to be necessary, if a subject is to think about himself self-consciously, that he actually have any information about himself. A subject may be amnesiac and anaesthetized, and his senses may be prevented from functioning; yet he may still be able to think about himself, wondering, for example, why he is not receiving information in the usual ways. But it would be (…) wrong to conclude from this that self-consciousness can be explained without reference to the various ways that subjects have of gaining knowledge about themselves (…). It is essential, if a subject is to be thinking about

instantaneously being able to locate it. Still, if the physician would ask us to locate this pain, we would understand his question. Similarly, we may not always be able to react to a pain, but if someone asks us what we wanted to do when we were inhibited from reacting to the pain, we would typically have an answer to this.

Evans rightfully observed that this knowledge that informs our IEM-judgments shows that IEM-judgments are not made in a solipsistic mind, as some may, at first, have thought.

> "(…) our thoughts about ourselves are about objects – elements of reality. (…) we are able to conceive of endless states of affairs involving ourselves, and what we conceive is not necessarily what it is like for us, or what it will be like for us, to be aware of, or be in a position to know the existence of, such a state of affairs. Therefore we are not Idealists about ourselves, and this means that we can and must think of ourselves as elements of the objective order. All the peculiarities we have noticed about 'I"-thoughts are consistent with,

---

himself self-consciously, that he be disposed to have such thinking controlled by information which may become available to him in each of the relevant ways." (Evans 1982, pp.215- 216)

Also: "I have emphasized that a subject's Idea of himself does not require him to have a current conception of himself; what is required, in the exceptional circumstances in which the various avenues of self-knowledge are blocked, is that the subject be disposed to accept any information accessible in those ways as germane to the thoughts we regard as manifesting self-consciousness. But in the normal situation, of course, these dispositions are exercised, and he has an evolving conception of himself, embodying information derived in the various ways, and partly retained in memory, which informs his thoughts about himself. As with other thoughts which are information-based, there is a presupposition that there is just one thing from which the various elements of the conception derive." (Evans 1982, p. 249)

> and, indeed, at points encourage, the idea that there is a living human being which those thoughts concern" (Evans 1982, p. 256).

Evans concludes that because "many philosophers give the quite mistaken impression that it is only our knowledge of our satisfaction of mental properties which gives rise to judgments exhibiting immunity to error through misidentification" (Ibid, p. 217), they get "the impression that in thinking of oneself self-consciously, one is paradigmatically thinking about oneself as the bearer of mental properties, or as a mind – so that our 'I'- thoughts leave it open, as a possibility, that we are perhaps nothing but a mind" (Ibid). He argues that if we would see that "our self-conscious thoughts about ourselves also rest upon various ways we have of gaining knowledge of ourselves as physical things" (Ibid, p. 213), we would also see that "[o]ur thoughts about ourselves are in no way hospitable to Cartesianism. Our customary use of 'I' simply spans the gap between the mental and the physical, and is no more intimately connected with one aspect of our self-conception than the other" (Ibid, p. 256).

While Sydney Shoemaker contributed significantly to our understanding of self-reference, Evans better covered, distinct occasions at which this self-reference is also an expression of self-knowledge that is immune to error through misidentification. Evans also elaborated more extensively upon the nature of this self-knowledge. Shoemaker taught us that we can refer to ourselves, without thereby envisioning properties which characterize us, and that this self-reference is immune to error through misidentification. Evans additionally clarified that it is not because this self-reference is without criteria, that it is not made possible in the first place by previously-acquired knowledge about ourselves and the world in which we live.

He also added that part of this knowledge may be knowledge about ourselves as physical objects, and that there are cases, in which our self-reference is immune to error through misidentification, even when we refer to ourselves as embodied subjects. This taught us that a self-conscious thought, of which, at first sight, we may think that it is about our inner consciousness and only reached by consulting this consciousness, in fact, can be about ourselves as embodied subjects, and only come about because we gained knowledge about, among other things, being a physical object in a physical world.

2. Some references that are IEM seem to refer to a mental self, but, in fact, refer to a physical self

Apart from the just discussed self-references which obviously refer to an embodied self, there are also self-references which are immune to error through misidentification and which, at first sight, may seem to refer to a mental self, but in fact refer to an embodied self. The self-reference in memories is one such reference. I will demonstrate this through an examination of Shoemaker's analysis of the self-reference in memory.

With regard to memories, Shoemaker holds that "one does not use bodily identity as a criterion of personal identity in making memory statements about one's own past" (Shoemaker 1963, p. 207). He argues that this is so "not because one uses something else as a criterion but rather because one uses no criterion at all" (Ibid). When we refer to ourselves in memory, we do not first check whether certain elements

are in place, to then conclude that we must be the person we have memories about.[57] We immediately assume that the latter is the case. In Shoemaker's words:

> "What 'shows' a person that he is identical with someone who existed in the past is not, on this view, his consciousness of some fact, or set of facts, that is criteria evidence of his identity with that person; it is consciousness of that identity itself" (Shoemaker 1963, p. 39). Or again: "If what he remembers is that *he* broke the front window, then for him the question 'Am I the person who broke the front window?' cannot arise" (Ibid, p. 135).

This is also why Shoemaker takes our self-reference in memory to be immune to error through misidentification: according to Shoemaker the self-reference in memory is immune to error through misidentification, because it does not involve an identification in the first place.

Still, Shoemaker's analysis of memory can be seen as a first step in discovering that our self-reference in memory is a reference to an embodied self, even when this reference is immune to error through misidentification. Shoemaker argues that "no one can imagine remembering, or knows 'what it would be like' to remember … a migratory past" (Shoemaker 1963, p. 208). To this, he adds that he is "doubtful whether it makes sense to say even that one can imagine *seeming* to remember, or having *apparent* memories of, such a past" (Ibid). For,

---

[57] "I shall argue that while one does not use bodily identity as a criterion of personal identity when one says on the basis of memory that one did a certain thing in the past, this is not because one uses something else (some nonphysical fact) as a criterion, but is rather because one uses no criterion of identity at all" (Shoemaker 1963, p. 124).

> "[i]f one thinks that one can imagine remembering (or seeming to remember) having actually been in Paris, New York, and Moscow in rapid succession, one is perhaps thinking that remembering this would be something like remembering such a series of pictures. But a memory of a glimpse of what looked like the Kremlin cannot be called a memory, not even an apparent memory, of seeing the Kremlin unless the other memories associated with it, e.g., the memories of what happened before and after the glimpse, are compatible with the proposition that one was in Moscow at the time at which the remembered glimpse occurred or is remembered as occurring" (Ibid).

In our world, these memories will only appear as compatible when it is physically possible that I have been at the places, represented by the distinct memories in the time-span, during which the remembered experiences took place.

So, even though Shoemaker explicitly denied that, in memory, we refer to ourselves as traceable physical objects and reasoned that, since our memories immediately present themselves as being about our own past lives, we do not conclude this by checking whether we were really physically present at the event that we seem to remember, he simultaneously admits that for our memories to function as they do, in practice, we cannot have a migratory past. Specifically, this means that we cannot just switch bodies all the time and so be at one place in one moment and at a totally different place in the next. Shoemaker argues for this by saying that our memories should fit in a coherent context. If I remember having seen the Kremlin, it should also fit in my life-story that I saw the red square and once flew to Moscow. Again, if we were not embodied and did not live in the world in which we factually live, this required coherency may have consisted in something else. Maybe in another kind of life, we could have been at the Kremlin in one second and on Table Mountain in South-Africa in the next. If we

were not embodied, and we seemed to have been at Table Mountain seconds after we were at the Kremlin, this would make sense and be coherent. The only reason for Shoemaker to assume that we will only perceive memories as memories, if they are coherent in the sense that they do not reflect a migratory past, is his presupposition that, factually, we live in this world through one body.

When we take this into account, it appears that we can have a memory of something we did, without, therefore, first having to examine whether we were physically present at that occasion, all the while the perception of this memory, as a memory, still depends on our assumption that we are embodied in just one body. If experiential memory does not allow for a migratory past, every instance of experiential memory should be of an event that took place at one particular moment, at one particular place, in one particular life.

V. OUR OBJECTIVE PHYSICAL PLACES GIVES US A SUBJECTIVE POINT OF VIEW

In the previous paragraphs, I highlighted some ways in which our bodies, perhaps unexpectedly, have a part in how we identify and refer to ourselves. Before closing this chapter, it is worthwhile to mention a more obvious factor that can make us identify and re-identify selves, through their bodies, even when we judge that their consciousness is what is most essential, or important to them. It is the element that P.F. Strawson chooses not to touch upon in his *Individuals* (1959): our body

determines our point of view and thereby makes us so localize ourselves both as a physical object and as a conscious subject.

Whenever we perceive something, we perceive it from a specific point of view. We see and hear something from a particular angle and distance. Which point of view we perceive things from depends on the localization and orientation of our body. So, our body determines the content of our conscious experience. It fixes what we perceive of our object of perception. It also constitutes our awareness of the shape we have and the space we take up. For instance, the place where I am sitting and the direction my face is turned determines which part, and how much, I see of the hallway. I notice that when I walk around, I can see more of the hallway and that when I turn my face away, I see less of it. This makes me aware of both the fact that, in sitting somewhere, I do not sit everywhere (I take up a limited place in space), and of the fact that I have a backside, from which I perceive nothing.

Naomi Eilan has argued that this limited point of view, which we have at any moment, and which makes us realize that we are also physical objects which are localized at one particular place in space, at one particular time, is also what makes us localize ourselves over time:

> "when we have such a primitive theory of self location in play, we can draw a distinction between how things seem to the subject and how they objectively are, where how things seem to a subject over time just is the extended point of view, the subjective route" (Eilan 1995, p. 339).

Here, Eilan points out, that we judge that we have been at all those places from where we perceived the world. We always observed this world, from a specific angle, and so never saw everything in its entirety, yet, at the same time, we know that, objectively speaking, there are things in front of us that we do not see, such as the back of

the table or the rest of the hallway. In judging that we have been at those places, from where we have observed the world, we both judge that we were where we were conscious and that we were where we were physically speaking; we were at the places where our physical positioning made us observe the world from a particular point of view. This is one other reason that can make us judge that, *as conscious beings*, we were where our bodies were; the placing and mobility of our bodies makes us simultaneously aware of the fact that we have a subjective perspective on the world and that this perspective is determined by our objective place in an intersubjectively shared space and time. There would be no subjective viewpoints, if we did not also take up one particular place in the world of solid objects. Our bodies are, thus, constitutive of our subjectivity and determine where we are, and have been, as subjects.

VI. Conclusion

The aim of the last three chapters was to determine some of the factors that explain why our body can have the role it has in our ascription of numerical identity to selves.

I recognized that it may seem strange that we identify selves in this way, given that we often identify ourselves more with the way we think, than with the particular body we inhabit. Our ability to be conscious of our consciousness and the specific contents of our consciousness oftentimes seems more essential to us than the body that produces this consciousness. It seems to say more about who we are as specific persons.

I then highlighted how our body can have a part in the identification or reference to mental selves. With Strawson, I demonstrated that material bodies function as basic particulars, which are needed to re-identify particulars, and that our own bodies can help us to re-identify particular persons, given that our bodily characteristics are ascribed to the same entity as our more mental characteristics are. Our bodily and mental characteristics are both ascribed to persons and not to respective bodies and mental entities. I also demonstrated how, often, when we think that we are conscious of ourselves as mere conscious beings, this consciousness is, in fact, informed by our knowledge that we are physical beings, as well: this is so when we experientially remember something we did or thought; when we ascribe sensations to ourselves or experience them in the particular way we do; when we have thoughts about our mental character; and even when we refer to ourselves, without explicitly identifying ourselves.

Not to forget the more obvious, I subsequently sketched how our point of view, which is always determined by our bodily positioning, equally makes it the case that in thinking about our subjective route through time, we also assume that, in fact, we were always, for the time of our existence, somewhere physically in an intersubjectively shared spatial world.

In the next and last chapter, I will discuss which image of the self we adhere to, when we identify a self through the particular body it inhabits.

CHAPTER 7:
WHICH IMAGE OF THE SELF CORRESPONDS TO OUR ASCRIPTION OF DIACHRONIC PERSONAL IDENTITY?

By now, I have shown that, in many circumstances, we identify someone as numerically identical as long as he has the same (in the sense of 'materially continuous') body. I have also elaborated on how the properties of a physical body allow for this. In this final chapter, I analyze which kind of self is constituted by our practice of identifying a self as numerically the same, as long as he lives through the same continuous body. I will argue that this self remains formal to an important extent, yet add that it is equally crucial for the constitution of our more substantial character and moral being.

I. A MINIMAL SENSE OF BEING A DIACHRONIC SELF

The self whom we consider to be numerically identical as long as he is materially continuous, is a formal self. This is a self which is not kept together by any intrinsic, meaningful, important and internally coherent characteristics. It is solely unified by an element that is external to it and that does not say anything about its specific character, i.e. the material continuity of a physical body. This self is formal in the double-sense (1) that our way of ascribing numerical identity to it is rule-based or formal (we ascribe it when material

continuity is in place) and (2) that the notion of this self is rather empty or formal (ascribing numerical identity to it does not require that we know anything about its character, and it being numerically identical does not guarantee anything about its character: it can develop itself in all kinds of directions). Let me give some instances of when this self appears, to clarify this.

There are moments we are aware of the fact that we have a past or a future, without simultaneously being aware of how we were, or will be then. These illustrate well that we can have a *minimal* or formal sense of being a *diachronic* self. One such example is when we wonder what will become of us. We then assume that we have one future, in which we will make some choices and will develop one particular life. We know that our life can only take a certain number of directions, because we can only be at one place at one time and because our life-span is limited. But this is as specific as it gets. We do not, and need not, know exactly how we will develop. That is exactly what we wonder about. Hence, it is not a concrete imagination of being a particular character that gives us the idea that we will be a specific self. We have the idea that we will continue to exist, but this idea of ourselves, as a future self, can be formal. It can remain rather empty: we do not have to imagine how we will be like, to imagine that we will still be there.

We have the same idea about other persons. For instance, we can wish the best for a baby and so demonstrate that we foresee that it will have one particular future, without binding this future to a particular

character.[58] Here, we do not wish a particular character the best, but a self – whichever character it may later develop.

Something similar holds for how we can think about our past. This is perhaps the clearest with amnesic patients, but we can all think this way. It may happen, that an amnesic patient starts answering the question about what she did earlier that day or that week, but then realizes that she cannot remember this. She then knows that she has a past while she does not recall anything particular about it. This, in fact, happens quite often: many Alzheimer patients know of their forgetting and are frustrated about it. An amnesic patient may also demonstrate that she assumes that she has a future by expressing the wish to start driving again or to die soon. She may also wish a granddaughter well on her travels – even if she forgets soon after this that her granddaughter is traveling at all. These examples rebut an existing misconception as, for instance, expressed by Revonsuo when he declares the following:

> "Amnesic patients who cannot form new memories are doomed to live in a permanent present moment. They have lost the awareness of self as a temporally continuous being who has traveled a long road from the past to the here and now, to this very moment of present conscious experience, and who will be heading towards the future" (2010, p. 137).

Amnesic patients may fail to remember what they did in the past, as well as fail to execute certain of their intentions, but still have and work with the idea that they have a past and future.

[58] Cfr. Breeur and Burms (2008, p. 141).

This minimal idea that we can have of being a self provides us with an alternative to the reductionist account of personal identity that Sartre (1996) and Parfit (1984) offer. It also provides an alternative to Daniel Dennett's (1998) theory of personal identity as a narrative construct.

Reductionists start from the idea that what we typically call an identity or self, in fact, constantly changes. They often add that these changes cannot indeterminately be kept together by an overarching consciousness or memory, because there are too many gaps within our consciousness, to do so. From this, they conclude that what we call a person is not one, but can be reduced to a stream of changing mental instances and characteristics.

Daniel Dennett equally observes that there is no real entity such as 'self,' in the world, but only disjointed streams of consciousness and changing character traits. As we saw in the previous chapter, Dennett reduced this self, in 'Where am I?' (1998), to a functional organism, which is, at no place other than, where its functioning takes place, i.e. where it is receiving and processing input and producing output. Dennett takes his analysis one step further in 'Why everyone is a novelist'[59] (1988). There, he points out that the reducibility of selves does not prevent us from narratively constructing them. According to Dennett, we get a sense of being one particular self by creating a story about ourselves that makes it seem as if different streams of consciousness still belong to one unified whole.

[59] This text was published 10 years after 'Where am I'. I refer to the 1998-version of 'Where am I,' but it was originally published in 1978.

I propose the alternative view that we can say of persons that they are diachronic selves, which can each be seen as a particular unity, even though they can develop themselves in all kinds of directions and they are not kept together by an uninterrupted consciousness. This self is not a specific coherent character, neither is it a narrative construct that turns something, perhaps at first sight incoherent, into something coherent. Rather, it is a unity, of which we imagine nothing, but that it has one past and one future only. We can refer to this self as a particular self, without knowing narrative particularities about the life of this self. It is the kind of self that – as Kripke (1972, p. 96) correctly pointed out – does not give or lose its proper name, depending on the characteristics he establishes, yet that we would not identify with a being that is made up of completely distinct and non-continuous material. Although our notion of this self is very minimal and rather empty, it is still meaningful – it refers to a diachronic unity, and has real effects, upon which I will now elaborate.

II. OUR MINIMAL IDEA OF BEING A DIACHRONIC SELF IS CONSTITUTIVE OF US AS MORE SUBSTANTIAL SELVES

The formality of the self I just described does not prevent it from functioning in our idea about, and attitudes towards, the more substantial self or character, which we equally take ourselves to be. On the contrary, its unity constitutes meaning. Below, I will first sketch how this formal self, as anchor of our identity, is presupposed in how we perceive and conceive of people as particular personalities or

characters, as well as in our interest in, and concern about, ourselves and others. In a second move, I show how our anchored identity also gives us freedom and flexibility in contemplating and shaping it. I conclude, by relating these points, and illustrate how the nature of the formal self explains the simultaneous feeling one can have of being both associated with, and dissociated from, a certain body or character.

1. The formal self, as anchor point of our identity, constitutes meaning and value

1.1 The intertwinement of us as a diachronically existing formal self and as a particular personality

Our existence as a particular personality, who is known to have a particular character, is not independent of our existence as a formal self, which remains merely numerically identical. It is intertwined with it and presupposes it. This becomes clear when we look at the causal links between our thoughts and behaviour, and at the way in which these links shape our character.

Neither (1) our numerical identity, nor (2) our character, are simply constituted by a chain of mental events that are connected to one another in a linear way, i.e. merely because mental event 1 causes mental event 2, which, in its turn, causes mental event 3 to occur. (1) To see why a chain of thoughts, which is connected because these thoughts diachronically caused one another, does not automatically

constitute one *numerically identical self*, consider the following. An expression of a provocative thought can cause someone to be provoked, which, in its turn, can cause someone to be scared of this provoked person. We then have three causally linked mental events (a provocative thought caused the feeling of being provoked, which then caused a feeling of fear), all the while the provoker, provoked and scared are not the same person. (2) Neither are merely causally linked mental events sufficiently interrelated to constitute the *character* of a particular person. The fact that a particular mental event causes another mental event is not constitutive of a personality, in as far as it is part of a chain with linear causal connections between its mental thought-buckles. A causal link between mental events only becomes part of the constitution of a person with a particular character, when, and, in as far as, the past of a particular person makes it understandable that this causal link is there, as well as when, and, in as far as, the occurrence of this causal link does not just describe a relation between two mental events, but says something about the person, in whose life the first mental event (and his past history) caused the second thought or established the second behaviour.[60] Say, for instance, that a

[60] Cfr. Richard Wollheim (1980): "Mental connectedness is not just any two-term relation between mental events that belong to one and the same person. Mental connectedness is that, but it also satisfies the two following conditions. One: the earlier of the two mental events that it relates causes the later of the two events. And two: the later mental event is caused by the earlier mental event in such a way that it – the later event, that is – then passes on to the whole person the causal influence of the earlier mental event" (p. 304). "…the underlying structure is that of a three-term relation, relating mental event, mental event, and the psychology of the person" (ibid, p. 305) "…it is this second occurrence of causality, the onward transmission of a causal influence once generated, that is simultaneously the key feature of mental connectedness as far as personal identity is concerned and that feature of mental connectedness which it advocates have failed to recognize" (Ibid, p. 305).

particular realization causes one to cry and that this crying causes someone to think that he should do something to dispel the reason for crying. Then the mere causal links between the respectively first and second, and second and third mental event do not yet constitute a particular person or character. They are only part of such a constitution when the past of a person can say something about why this realization could make this person cry, and, in as far as, a person's willingness to do something about the reason for crying does not just say that crying caused this eagerness, but also characterizes the person with this will as, for instance, resilient.

This demonstrates the dependence of our existence as a particular character on our existence as a more formal, merely diachronically existing self. Phenomena such as the causal link between mental events are only indicative of a certain character, when they take place in one person's life, and when one can interpret their occurrence, in the light of this person's past, as well as assess what it additionally teaches us about the character of this person. So, the mere idea that a person has a past and future (and, hence, an extended numerical identity), is one of the conditions to interpret mental events as part of and constitutive of a character. Different character traits or causally related thoughts do not first constitute our idea of the diachronic existence of a self on merely other grounds. We must have the idea that we have a past (whatever its character), to turn to this past to interpret mental events as being part of, and further constitutive of a particular character. Of course, this past will mostly have a particular character that will help us see how it is exactly related to other events in a person's life. Yet, if we did not already have the idea that a person

has *a* particular past, we would not start to relate events, in this way and, even when we did, we would not know, to which character we should turn first, to explain the next causally related event. Different characters might explain it equally well. Our assumption of having one particular past and future is (at least) one factor which makes that we interpret characters, by means of the mentioned particular causal links.

1.2 Being interested in, and concerned for ourselves and others

Our (idea of our) existence as formal diachronic selves also has a part in the constitution of our interest in, and concern for ourselves and others. This may seem strange, at first. Our interest in others can seem to be merely generated by their peculiar personality, and our concern for ourselves may often seem to be a concern for one of our specific characteristics (we may want to change them or hope not to lose them). Still, there are aspects of the formality of our existence as formal selves, which constitute particular interests and concerns that we have about ourselves.

1.2.1 Our curiosity about what will happen to us

In the case of our interest in ourselves and others, these aspects are that we expect that we all have one diachronic existence, while we also know that this does not determine which kind of characteristics we

will develop; we could develop different ones, but we will only develop particular ones. Together, this stimulates our curiosity about ourselves and others: just as our curiosity about the ending of a particular movie or book is aroused because we know that books and movies could have a whole range of different endings, although they, in fact, only have one ending, so we are curious about what may become of us or others because we realize that we could all develop ourselves in many different ways, while we also know that, in the end, we will each just have lived one particular life.[61] Here, our interest in ourselves is not guided by any particular imagination of how we will be like. We are just interested in what we will be like, because we assume that we will still exist (in one – and only one – shape or another).

1.2.2 Our concern about what will happen to us

A reference to ourselves as formal selves is also needed to describe particular concerns that we have about ourselves. I referred to one such concern in my discussion of Parfit in the first chapter. Parfit (1984, pp. 215, 216, 217, 224-225, 241, 245, 255, 273-275; 1995 29, 33, 34, 44; 2007 passim) argued that when we seem to be concerned about our own future or survival, we are solely concerned about ourselves as qualitative beings, who can realize certain projects, thanks to these qualities. From this, Parfit concluded that our apparent concern for

[61] This precise example has been given by Arnold Burms at several of his lectures at the KU Leuven (2006-2012)

our own future or survival is not really a concern for *our* future or survival. He reasons that, as soon as we realize that we care about ourselves as qualitative beings, who can accomplish some projects, we should admit that it is immaterial to us whether we survive, or whether someone else with our qualities finishes off the projects we set up.

Against Parfit, I argued, with Williams (cfr. chapter 1), that our concern for our own future can be about us as merely numerically identical beings (i.e. minimal, formal selves), rather than as replaceable qualitative beings: even when one tells us that we would lose our memories and change our character, we would still worry about what would happen to us after this (Williams 1973b). This is, for example, what happens when one expresses the wish to no longer live after a certain stage of Alzheimer's. Though, in that case, we will have forgotten who we were before, and, even though we will have changed our character, this is not a wish for murder. It is the wish to then die oneself. I can, thus, be concerned for a self, of whom I formally determine that this will be me, but to whom I further do not bear any connection.

1.2.3 Our desire to accomplish projects ourselves

Perry (1976) offers another good counterargument against Parfit's vision: we do not only have the wish to accomplish our projects, but also to personally accomplish them. Perry hypothesizes that the importance that we here attach to identity is derivative. He suggests that we came to attach this importance to identity because it proved to

be evolutionary beneficial to do so; human kind found that persons tend to get the return of investment of what they personally do. I support this hypothesis, as long as the satisfaction and pleasure we can derive from how others recognize, esteem, and remember us are taken up as elements in the return of investment. However, this derivativeness of the importance of identity does not make it so that we can, here, reduce the object of our concern to someone or something, of whom the numerical identity is immaterial to us. While the importance that we attribute to this identity may be motivated by other factors than this identity itself (such as its evolutionary benefit or the pleasurable feeling of being esteemed), it is still us, qua numerically identical being, whom we are concerned about, and not someone who, say, has similar feelings of pleasure as we do. So, there is an aspect to our concern for ourselves that is a concern for ourselves qua formal selves, i.e. qua merely numerically identical selves.

1.2.4 An incentive to shape ourselves as liable (moral) characters

A last aspect of our formal self that has an effect on our concern about ourselves, is the fact that its physical existence is never interrupted. Our diachronic identity is bound up in our physical continuation. This means that we never go temporarily out of existence and that we are always observable. The impossibility for us to temporarily disappear can enhance[62] our concern for what we

[62] I say 'enhance' because the body, alone, does not create this concern. I can only have this concern when I am also gifted with a self-consciousness that is

should try to be like. If we were not always there, or were not always observable, we would not have to be concerned about what we are like or how we appear to others. We could choose to retreat from being any particular person, for some time, and be happy with that prospect. But given that we are always identified with whichever personality is living through one particular material body, and will always be remembered as this personality, most of us are concerned about this personality and try to shape it into one that will be appreciated and one that we would like to be remembered as. This is often a reliable, and so, more or less coherent one. So, our assumption of, among other things, being a diachronic material entity has an influence on a concern that we have about ourselves, and on the shaping of our (moral) character.

Just like this, the realization that one has a limited lifespan and can only be at one place, at one time, has an influence on the character one develops. It makes one realize that one will have to make choices about where one wants to be and in what one wants to engage oneself.

---

sophisticated enough to allow me to think of myself as a diachronically existing being, who has a certain capacity to determine how he is, who appears in a certain way to others, and who can become a more or less desirable subject. The body can enhance this concern because it really confronts me with the fact that I can never temporarily disappear, as well as that my history stretches further than my memory does, and that I can even be held accountable, and so feel accountable, for what I do not remember.

1.3 Promise and desire

In the previous paragraphs, we saw how the formal self serves as an anchor point that allows for the development of a character, as well as it can make us esteem a particular life, in a particular way, i.e. generate our particular curiosity and concern for it. It makes us concerned for one particular life rather than another; want to finish projects *ourselves*; and feel responsible for what we do during our lifetime. I will now elaborate on the phenomenon of promising, another aspect of our moral behaviour of which the formal self, as anchor point of our identity, is constitutive.

When we promise something, we are expected to keep this promise, even when our desires change. That is the nature of a promise: to bind yourself to do something and not invoke excuses not to do so. When someone promises us something, we expect *him* to keep *his* promise, regardless of how his desires or character change. So, in the promise, we assume that someone can stay the same, regardless of changes in his character or desires. We thus work with the idea of being formal selves; i.e. selves who have a diachronic identity that does not depend upon a specific set of character traits.

This presupposes the formal self that I have described, and distinguishes my view, once more, from that of Sartre (1961, 1996, 2005). Just like Sartre, I have recognized that our consciousness is free. We cannot guard it with a particular personality. All kinds of thoughts may occur to us. However, contrary to Sartre, I do not conclude from this that we are no particular ego. While Sartre argues that your current mental state and actions constantly redefine yourself and that

you can never truthfully be someone in particular just because of your past actions, I hold that past actions do constitute our history and identity and argue that this is, for instance, shown in how we can be held accountable for our promises. Without the formal self I described, our concept of a promise could not exist.

What holds for promises does not hold for desires. We are not typically bound to them. Still, as became clear in the previous paragraph about promise, we can remain the same person, even if our desires change. When a later desire conflicts with a previous one, there are typically two choices. (a) When the first desire did not place me under any moral or legal obligation by promising something to someone or signing a contract, my later desires will most typically overrule my earlier desires in their quest for being respected.[63] Say, for instance, that someone, at one point in his life, expresses the wish to have euthanasia performed on him, if he were ever to suffer from a specific medical condition, but once in this condition, revokes this desire. Then, we would not hold him to his previous desire. This is why in Belgium, where euthanasia is allowed if the patient suffers unbearably, the patient is asked to restate his euthanasia desire twice, if his condition still allows him to do so. (b) One can feel justified in accommodating a previous desire rather than a later one, if the desiring subject, at one point, expressed that he considers this previous desire to be of a higher order than a potential contradicting later one. So, someone can feel justified in not providing a man with a cigarette when

[63] For the same statement in relation to the noble man case of Derek Parfit, which I will criticize shortly, see Christin Korsgaard in Marshall 2012.

asked for it, if this man said earlier that, even if he would at one point desire a cigarette and ask for it, he'd wish that others would not give it to him, thereby helping him to quit smoking. Our respect for respectively the later or earlier desire in case (a) and (b) does not stem from our consideration that the contradicting desires come from distinct persons and that one of these persons, for some reason, deserves more respect than the other. Instead, we respect a hierarchy of a set of desires that one person can have throughout his life-span.

The previous explains why Parfit's conclusion in his famous nobleman case is drawn too quickly. Parfit (1984, pp. 327-329) imagines a Russian nobleman who wishes to leave his entire inheritance to the peasants, but is afraid that he might become less of a socialist later on in life and then might revoke this wish. To prevent this from happening, he makes his wife promise that she will, in this case, not listen to his later self, but instead respect the wish he currently has. He considers this wish as essential to him and tries to convince her that he would no longer be the same person without this wish. When the wife acts accordingly, Parfit concludes that she does this because she loves her late husband who expressed this wish, and because the remaining man with another desire is not, in fact, the nobleman, but his inheritor who is not in a position to decide what happens with his inheritance.

Parfit correctly judges that if we feel that, not the later, but the earlier man's wish should be respected, we, in fact, consider the later man as the inheritor of this earlier man, rather than as this man himself. Ignoring the later man's wish is only legitimate in this case. As we saw, we would not let the earlier wish overrule the later one, if

the man were still the same, for this is not an addiction case, in which a desired hierarchy of desires would apply. However, Parfit is too quick to conclude that, if someone, at one point, changes a desire he considered to be essential to him, this means that he dies and then merely becomes his own inheritor. He could change this desire without losing his identity. As we saw, we would, in this case, mostly consider the later desire as the one to be respected, except if he is under a moral or legal obligation that forbids us to do so. We would also make an exception, if his first desire functions as a higher order desire over his later desire.

Apart from this, Parfit's treatment of the noble man case contradicts his own treatise on personal identity. In his section on personal identity in *Reasons and Persons*, Parfit (1984) argues that we should rid ourselves of all concerns related to personal identity, given that there are no entities with such guaranteed identity, but just gradually ever changing beings. Yet, the noble man fully identifies with one image of himself – a generous socialist – and has worries, which he would not have if he did not think of himself as being one such specific character. To accept this worry as legitimate and use it as a starting point for one of his arguments goes against for what Parfit otherwise argues.

Contrary to what Parfit actually concludes about the noble man, but in line with what he otherwise contends, I argue that there is an illusion to how Parfit's noble man thinks about his identity. It is wrong of him to identify with one image of himself. As Parfit contends, I think he is subject to changes. However, from this, I do not infer, as Parfit does, that all ideas of being a particular identity are illusionary. I draw

a distinction, which Parfit does not draw, between imaginary identification which is illusionary and formal identification which corresponds to a reality in the intersubjective world in which we factually live. It is illusionary to identify yourself with one image of yourself, because you can, in fact, change and still remain the same. Yet, you are justified in having the feeling that you are one diachronic identity with one history and one future, all the while not knowing more about the character of this entity. These two forms of identification are radically different. In the first case, we can explicate exactly what we think our identity consists of; in the second case, there is not so much to say about it, because the identification remains formal. Further, the first form of identification does not allow for considerable or radical changes in character, while the second does.

With what went before, I hope to have shown that the effects of the formal self as the anchor point of our identity, and, so, the formal self, are real. In the next section, I will shift attention to the flexibility of the development of our identity, for which this formal self allows.

2. The formal self gives us freedom in thinking about and effectively changing ourselves

Apart from functioning as an anchor point of our identity, or more precisely, because it functions as an anchor point of our identity, our formal self also gives us freedom and flexibility in thinking about and effectively changing ourselves. The minimal idea of being a particular, numerically identical entity, about which nothing definitive is known,

except that it is diachronic, is presupposed in the thought about and effective changing of ourselves. We can dream about becoming different – not just slightly, but radically. I can, for instance, think about how great it would be to get rid of my laziness and food cravings; to be athletic, tall and thin, instead of small and chubby; and to be witty and social, rather than shy and inhibited. I then think that this would be great, because *I* would be like this, not because someone else would be like this. This 'I' is a minimal self. I have no other idea about it, except that I would be it. I do not tie it to any specific set of character traits.

One could object that I do not, in this fantasy, really dream about acquiring different characteristics, but about becoming another person. Or one could object that, even if I do not *dream* about becoming another person in this fantasy, I would *in fact* cease to exist and another person would be born, if this fantasy materialized. This would undermine my statement that a formal self is involved, here, and is what makes this phenomenon possible. Against this conclusion, I argue that the two latter phenomena differ radically from the one I described. An elaboration upon this difference will not only confirm that a minimal self is presupposed in the imagination I first described. It will also shed more light upon the nature of this formal self.

To counter the first objection and to distinguish fantasizing about being a different person from imagining acquiring a radically different character yourself, it is necessary to first determine what it means to fantasize about being a different person. This can mean different things, depending on which kind of fantasy you embark upon. I will

look at three. The first is that of a child watching a television-series, screaming 'that's me,' and then starting to act as this particular character. The second is that of an adult who imagines what he would do, if he were someone else. The third is that of someone who imagines being someone else and experiencing everything from his perspective. I will demonstrate that, in none of these cases, we imagine to really be someone else. This demonstration will highlight the differences between imagining being someone else and imagining changing drastically. The apparent impossibility of the former imagination will also take away the likeliness that imagining becoming endowed with different character traits amounts to imagining that one becomes someone else.

I will first argue that neither the child, nor the adult, who imagines what he would do if he were someone else, thinks of himself as really being this other character. The child just demonstrates, and the adult just thinks about, what he would do if he had certain characteristics of this other character and encountered a situation like they encounter. In doing so, the child plays a role, like actors do. He can do this with a lot of passion and even forget about himself, but he does not think that he really is this character. He just enjoys the imagined powers and environment, in which he places himself for a moment. Something similar holds for the adult who wonders what he would do instead of someone else. In distinction with the child, this adult does not exactly act like an actor in a play, but more like an actor preparing his role. He tries to imagine what the character would do in certain circumstances given his history, hopes for the future and surrounding people. To make this more realistic, the adult thinks about the choices he would

make in similar circumstances. These might be slightly different than those of the character, in whose shoes he tries to place himself. If our adult were confronted with a particular situation, his proper character traits might make him take different decisions: he could, for instance, be more averse to risk than the character about whom he fantasizes. This generates the difficulty we always have when someone asks 'what would you do if you were me'. We are then asked to give advice, with respect to another person's life, and so try to envision the specifics of his life, character, and desires for the future, as well as we can. At the same time, we are asked to say what *we* would do if we were him. In order to give a well-thought answer, we also ponder whether what we think is desirable: we consult our own inclinations. Imagining what you would do if you were someone else is thus not fantasizing about really becoming this person, but rather attempting to find out what you, an empirical character with certain developed responses, would do when placed in a particular situation, in which you would have a different history, set of relations, and also some diverging characteristics and desires. You remain aware of the fact that you are not the person to whom you try to give advice. You base your advice on your thoughts about his situation, not on an imagination of really being him.

We can, thus, discard of the previous two fantasies. Contrary to what may, at first, seem to be the case, they are not truly fantasies of being someone else. A third kind of imagination is more promising. This is the participatory imagination of being someone else, in which someone imagines that he experiences the world from this other person's perspective. This participatory imagination has been the

object of philosophical investigation for Bernard Williams (1973a), and Zeno Vendler (1984).

In his 'Imagination and the Self,' Williams demonstrates how, even in such participatory imagining, we do not truly imagine that we *are* someone else. Again, we are just role-playing. Although he does not explicitly say which kind of role-playing involves participatory imagining, Williams could, here, think of an empathic way of acting, in which actors do not just apply a particular technique to make the audience understand which emotion his character has, but in which the actor evokes those emotions in himself that he imagines he would feel if he were this personality, so as to credibly represent them. We can further think of this actor as forgetting about the public that watches him and just seeing the surrounding scenery as he imagines his character would, or must have seen it.

While role-*playing* does not amount to imagining really *being* someone else, we can show how this last kind of imagination, which we may, at times, possibly *call* an imagination of being someone else, differs radically from imagining that one's character drastically changes. In the latter case, a formal self is involved: in imagining how most, or all, of your character traits would change, you get an image of an attenuated self, which persists in spite of these changes. But there is no role for such a self in participatory imagining that you are someone else. For, as Williams explains, there are only two selves that have a function in this role-play: the empirical self that imagines, and the self of the character that he imagines being and that has a

particular view of the world.[64] When I imagine being this character, I imagine experiencing everything from their shoes and my imagination can be so lively that I actually sense what this character could sense. It is, however my empirical self that senses this: I hear what I hear through my auditory system, and sense what I sense through my nervous system and metabolism. I am not, nor imagine that I am the same attenuated self, now, with different characteristics, as the character I play.[65]

Zeno Vendler understands this participatory imagination differently. According to Vendler (1984), we can imagine that we could really have been someone else. We can, for instance, imagine that we would have been born elsewhere, at another time, in another body, demonstrating other talents and inclinations. What would remain the same – and what is enough for Vendler to conclude that *I* could have been this person – is the framework that allows for something like human consciousness, i.e. the fact that the consciousness is both synchronically unified and diachronic. This allows for someone to be a self: his integrated experience can give him the sense of being at one particular place and of being one subject who has distinct sensations, just as his sense of temporality allows him to feel that he is one subject which had, has and probably will have experiences. Inspired by Kant,

---

[64] "… what I am doing, in fantasy, is something like playing the role of Napoleon" (Williams 1973a, p. 44). "In the description of this activity, only two people need figure: the real me and Napoleon. There is no place for a third item, the Cartesian 'I', regarding which I imagine that it might have belonged to Napoleon" (Williams 1973a, p. 44).

[65] " … the fact that I can, in the only way that arouses my interest, imagine being Napoleon has no tendency at all to show that I can conceive, as a logical possibility, that I might have been Napoleon" (Williams 1973a, p. 45)

Vendler calls the subject who can have this kind of unified and diachronic experience 'the transcendental self'. He then argues that I would have been this person if my transcendental self were his transcendental self. Vendler concludes this from the following reasoning.

> "…in performing transference, i.e. in imagining being someone else, the object of the exercise cannot be fancying *me* (Z.V.) being identical with that person. In imagining, for instance, being Ronald Reagan, I cannot be imagining the identity of Z.V. with R.R., for it is patently impossible for these two men to be one and the same, and the patently impossible cannot be imagined. Therefore … being R.R. must simply consist in fancying being in a certain possible state, without involving Z.V. in the object of the exercise" (Vendler 1984, p. 105).

But what, then, with the 'I' who still seems to have a function in this imagination?

> "In imagining being R.R. I conjure up something which is not the case. What is not the case? That I am in fact R.R. Now this 'I' cannot mean Z.V., since in doing this transference I do not imagine Z.V. to be R.R. Therefore, it seems, 'I', the subject of this act, is not identical with Z.V." (Vendler 1984, Ibid).

It is my transcendental self:

> "…nothing will be left beyond the 'conditions all experience preserving the 'format' of the unity of consciousness" (Vendler 1984, p. 106-107). "[This] transcendental 'I' is not a thing, it has no content; it is a frame into which any content may fit, or, better, a 'format' or a 'schema' into which any content may be cast" (Vendler 1984, p. 107).

Vendler's conclusion is false and puzzling. It is false because the phenomenon of the transcendental self cannot serve as an identifying criterion for particular persons and, hence, not determine that *I* would

be a particular personality; i.e. that it would really be me who is a particular personality. It is puzzling because Vendler admits this, and so highlights a remark that Kant – from whom he got the idea of the transcendental self – already made.

> "The identity of the consciousness of Myself in different times…is only a formal condition of my thoughts and their connection, but it does not prove at all the numerical identity of my subject, in which – despite the logical identity of the I – a change can go on that does not allow it to keep its identity…" (Kant 1998, A363) "[The] identity of person in no way follows from the identity of the identity of the I in the consciousness of all the time in which I cognize myself" (Kant 1998, A365).

Hence, being a transcendental self can never determine that I am a particular personality, such as Ronald Reagan, as Vendler strangely admits himself:

> " 'Am I entitled, then, to my own little transcendental self, distinct from yours, which has the same kind of access to my mind as yours does to yours?' No, you are not, I reply, but nor am I. I have a mind, but I do not 'have' a transcendental self: the ultimate subject cannot be had….how, in the first place, could there be many *contentless* beings, not located in space and time either? How would they differ from one another? 'By the relations they bear to this and that individual mind', you suggest. If so, I reply, then 'my' transcendental self would be essentially tied to me, thus, once more, transference would be impossible: I could not imagine Ronald Reagan, because I could not *be* in the state of being he. But I can" (Vendler 1984, p. 109).

Moreover, it cannot even determine that I am, or remain, a particular 'I'.

> "For we cannot judge even from our own consciousness whether as soul we are persisting or not, because we ascribe to our identical Self only that of which we are conscious; and so we must necessarily judge that we are the very same in the whole of that time of which

> we are conscious. But from the standpoint of someone else we cannot declare this to be valid because, since in the soul we encounter no persisting appearance other than the representation 'I' which accompanies and connects all of them, we can never make out whether this I (a mere thought) does not flow as well as all the other thoughts that are linked to one another through it" (Kant 1984, A364).

It is with the transcendental 'I,' as with the feeling of mineness, to which I referred before. It is a function of everyone's consciousness, and gives one the feeling of being an 'I', but the personal experience of being a self, which we get from this, does not render us numerically identical with one specific substance, over another.

With this, yet another interpretation of what it could mean to imagine being someone else appears not, in fact, to refer to an imagination of *really* being someone else. We have now seen many interpretations of what imagining being someone else could amount to. None were able to demonstrate that I can conceive, as a logical possibility that I might really have been or become someone else.[66] A fortiori, none can prove that 'imagining to change drastically' is in fact 'imagining to be someone else'.

Yet, even when the phrase of 'imagining to be someone else' does not exactly run the gamut, the previous analysis of what it can mean to imagine to be someone else tells us something about the distinction between this phenomenon and imagining to change drastically in

[66] Neither does our imagination of being someone else need to have this tendency: "the fact that I can, in the only way that arouses my interest, imagine being Napoleon has no tendency at all to show that I can conceive, as a logical possibility, that I might have been Napoleon" (Williams 1973a, p. 45). I can just enjoy thinking about what to do in certain scenarios, as well as about what it would feel like to be put in such a situation.

character. In the latter case, *one* self is presupposed: you imagine that *you* will get other character traits; i.e. that you would be very different, but still numerically the same. As we saw, no such self is presupposed in imagining being someone else. It is not required. I can just role-play as an empirical self and think about an empirical self who existed or could exist. Moreover, such a numerically identical self *can* neither be there in the form of an empirical self, nor in the form of a transcendental self. An empirical self, born in a certain place, at a certain time, and in a certain body cannot be numerically identical with an empirical self, born at another place, time, or in another body. A transcendental self, in its turn, is not a substance to which certain (re-)identification rules apply.

So, the distinction between these two kinds of imagination tells us more about the formal self, with whom we identify when we either merely fantasize about how nice it would be to be attributed some different character traits, or when we really try to shape ourselves by changing some of our most persisting characteristics. This formal self is tied to a more substantial self: it can be attributed many different character traits, but never really live the life of someone else. While this attenuated self is rather empty and formal, it is not a current framework for consciousness as the transcendental self is. Instead, it has one particular history, which is the history of a particular person and coincides with his physical history. It could not have come into existence at another place and time than it actually has.

There's another reflection on imagination that can help us understand how our idea of the formal self functions. It is that of the

distinction between conceiving and visualizing and was, for instance, drawn by Bernard Williams (1973) in his 'Imagination and the Self'. There, Williams addresses the question whether one could imagine an unseen tree. One might be reluctant to admit that this is possible and argue that to imagine an unseen tree is to visualize it, which makes it into a seen tree, rather than an unseen one. Williams demonstrates how this thought is invalid. I can conceive of something without visualizing it, and even when I visualize it in imagination, what is visualized is not necessarily the imagined.[67] So, we can conceive of an unseen tree without visualizing it, and just think of the possibility that there is an entity somewhere of which we do not know the exact shape, but which can be called 'tree,' and is unseen by anyone. In doing so, pictures of a tree in a desert may come to mind, but this does not imply that this visualized tree is also what I am really imagining. I can be aware of the fact that the trees I currently visualize are not the unseen tree, of which I conceive the possibility. Besides, my visualization of a tree in my imagination does not make it seen: "in visualization nothing is *really* seen"[68] (Williams 1973, p. 37, my italics).

In working with the idea of the formal self, we equally conceive of, rather than visualize it. This allows us not to tie the formal self to any of the specific characteristics that are always represented in

[67] "… for certain purposes at least, and for certain applications of 'imagine', we can properly make the determinant of *what he imagined* his imaginative project, and not what he visualized, if the visualized anything" (Williams 1973a, p. 33).

[68] "…even if visualizing is in some sense thinking of myself seeing, and what is visualized is presented as it were from a perceptual point of view, there can be no reason at all for insisting that that point of view is of one *within* the world of what is visualized; any more than our view of Othello is a view had by one in Othello's context, or the cinematic point of view is necessarily that of one stealing around the characters" (Williams 1973a, p. 37).

visualization. It also offers an alternative for an impersonal account of personal identity. Someone may claim that you should not identify yourself with a current image of yourself, because you will gradually evolve. He may then conclude that you should not identify yourself with any entity, at all, because you never stay exactly like this entity. If the formal self is a self of which we *conceive* that it remains numerically identical, the previous is not a necessary conclusion: my conception of myself as one numerical entity, is then, in any case, not tied to any specific image of what this entity should resemble.

III. THE ELUSIVENESS OF THE SELF

Together, this formality of the self, about which I spoke in the first section of this chapter, as well as its both anchoring and flexibility allowing function, which I just described, can give us a new explanation of the elusiveness of the self. Who or what we really are often seems to escape us. Some maintain that we can get rid of this elusiveness by discovering who we really are, i.e. by discovering what our inner essence is, which character we really have, what we really want. Some philosophers choose to explain this elusiveness by arguing that there is no self, i.e. that it is just a fiction (Hume 1978), a to other relations reducible entity (Parfit 1984), or a narrative construct (Dennett 1988). Contrary to this, my understanding is that we experience the self as elusive because the consciousness-producing continuous body, which is crucial for its constitution, makes that, while it can be followed, it can never be fully appropriated. The continuous

body allows us to follow a self through space and time. It makes it easier for us to see a character develop, and hold someone responsible. At the same time, the body does not tell us much about who we are. It functions as the words in a poem. Just as a poem would not be without its words, and its meaning and beauty cannot be abstracted from them, all the while the letters of these words are not the content of the poem, so we can only discover a rich character, by following someone through his body, all the while this body as such neither is, says, nor guarantees much concerning this character. Moreover, just as a poem is, but can never be replaced by how someone interpreted it, so we cannot fully detach ourselves from what others see in us, while, at the same time, we never fully coincide with the existing images of us. We are just what we do and exhibit. There is no inner meaning without medium. At the same time, images that we, and others, have of us remain external to us and are snapshots that never bring together everything we are or will be.

The above also explains why we can sometimes feel discomfort when we are associated with a certain self. We are born into it and have no real say in how it is followed. Our existence, as diachronic numerically identical entities, allows us to develop a substantial personality that can be recognized and loved, but it is never a guarantee for any coherency or continuation of the qualities we further develop. We can be seen and numerically identified, while, at the same time, we can be uncertain about what the qualitative link is between how we are now, how we were, and how we will be.

BIBLIOGRAPHY

Alexander of APHRODISIAS, *De Anima*. Translated in: Harris, James 1765, *Hermes or a Philosophical Inquiry Concerning Universal Grammar. Second edition*. London. Quote found in: Copenhaver, Rebecca (April 2006) Thomas Reid's Theory of Memory, *History of Philosophy Quarterly*, Vol. 23, No. 2, pp. 171-188.

AYERS, M 1996, *Locke. Epistemology and Ontology*, Routledge, London.

BAKER, L R 2000, *Persons and Bodies. A Constitution View*, Cambridge University Press, Cambridge (UK).

BERNECKER, Sven (2007) Remembering without Knowing, *Australasian Journal of Philosophy*, Vol. 85, No. 1, pp. 137-156.

BREEUR, Roland and BURMS, Arnold (2008) Persons and Relics, *Ratio*, Vol. 21, No. 2, pp. 134-146.

BUTLER, J (1975) [originally published in 1736 as Appendix of Butler's *Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*], 'Of Personal Identity', in J Perry, *Personal Identity*, University of California Press, Berkeley, pp. 99-105.

CAMPBELL, C 1994, *Past, Space and Self*, MIT Press, Cambridge (MA).

CASSAM, Q 1997, *Self and World*. Oxford University Press, Oxford.

DENNETT Daniel C (1988) Why Everyone Is a Novelist, *The Times Literary Supplements*, No. 4459, pp. 466-474.

— 1998, 'Where Am I', in D Dennett, *Brainstorms: Philosophical Essays on Mind and Psychology*, MIT Press, Cambridge (MA), pp. 310-323.

DESCARTES, D 2004 [originally published in 1639], *Méditations Métaphysiques*, PUF, Paris.

EILAN, N 1995, 'Consciousness and the Self', in J L Bermúdez, A Marcel, N Eilan (eds), *The Body and the Self*, MIT Press, Cambridge (MA), pp. 337-357.

EVANS, Gareth 1982, *The Varieties of Reference*. Clarendon Press, Oxford.

GARRETT, D 1981, Hume's Self-Doubts about Personal Identity, *The Philosophical Review*, Vol. 90, No. 3, pp. 337-358.

GAZZANIGA, M.S. 1967, The Split Brain in Man, *Scientific American*, Vol. 217, No. 2, pp. 24-29.

GEACH, Peter (1967) Identity. *Review of Metaphysics*. Vol. 21, No. 1, pp. 3-12.

HUME, D 1978 [originally published in 1739], *A Treatise of Human Nature*, Oxford University Press, Oxford.

JUDSON, Lindsay (1987-1988) Russell on Memory, *Proceedings of the Aristotelian Society, New Series*, Vol. 88, pp. 65-82.

KANT, I 1998 [originally published in 1781], *Critique of Pure Reason*, translated and edited by P. Guyer and A.W. Wood, Cambridge University Press, Cambridge (UK).

KLAWONN, E 1987, The 'I'. On the Ontology of First Personal Identity, *Danish yearbook of philosophy*, Vol. 24, pp. 43-76.

KRIKPKE, S 1972, *Naming and Necessity*, Harvard University Press, Cambridge (MA).

LEIBNIZ, G W 1981 [originally published in 1704], *New Essays on Human Understanding*, translated and edited by P. Remnant and J. Bennett, Cambridge University Press, Cambridge (UK).

LOCKE, D 1971, *Memory*. Macmillan, London.

LOCKE, J 1996 [originally published in 1690], *An Essay Concerning Human Understanding*, abridged and edited by K. Winkler, Hackett Publishing Company, Indianapolis.

MARSHALL, R (2012) Treating People as Ends in Themselves, *3:AM Magazine*, www.3ammagazine.com, accessed June 30th 2013.

MARTIN, C.B., DEUTSCHER Max (Apr. 1966) Remembering, *The Philosophical Review*, Vol. 75, No. 2, pp. 161-196.

MCDOWELL, J 1997, 'Reductionism and the First Person', in J. Dancy (ed.) *Reading Parfit*, Blackwell, Oxford, pp. 230-250.

MENDUS, S 1980, The Two Analogies in Hume, *The Philosophical Quarterly*, Vol. 30, No. 118, pp. 61-68.

NIGRO, G and NEISSER U (1983) Point of View in Personal Memories, *Cognitive Psychology*, Vol. 15, pp. 467-482.

OLSON, E 1997, The Human Animal. Personal Identity without Psychology, Oxford University Press, Oxford.

O'SHAUGHNESSY (1989) The Sense of Touch, *Australasian Journal of Philosophy*, Vol. 67, pp. 37-58.

PARFIT, D 1984, *Reasons and Persons*, Clarendon Press, Oxford.

— 1995, 'The Unimportance of Identity', in H Harris (ed), *Identity*, Clarendon Press, Oxford, pp. 13-45.

— (1999) Experiences, Subjects and Conceptual Schemes, *Philosophical Topics*. Vol. 26, No. 1 and 2, pp. 217-70.

— 2007 *Is Personal Identity What Matters?*, The Ammonius Foundation, http://www.ammonius.org/assets/pdfs/ammoniusfinal.pdf, accessed April 13th 2010.

— (2012) We Are Not Human Beings, *Journal of Philosophy*, Vol. 87, No. 1, pp. 5-28.

PERRY J 1976, 'The Importance of Being Identical', in A. Oksenberg Rorty *The Identities of Persons*, University of California Press, Berkeley, pp. 67-90.

REID T 2010 [originally published in 1785] *Essays on the Intellectual Powers of Man*, edited by J Bennett, www.earlymoderntexts.com, accessed November 8th 2013.

REVONSUO, A 2010, Consciousness. The Science of Subjectivity, Psychology Press, Hove.

RUSSELL, B 2001 [originally published in 1912], *The Problems of Philosophy*, Oxford University Press, Oxford.

SARTRE, J-P 1961, *L'Existentialisme est un Humanisme*, Les Editions Nagel, Paris.

— 1996, La Transcendance de l'Ego : Esquisse d'une Description Phénoménologique, Vrin, Paris.

— 2005, L'Etre et le Néant. Essai d'Ontologie Phénoménologique, Gallimard, Paris.

SHOEMAKER, S. 1963, *Self-Knowledge and Self-Identity*, Cornell, Ithaca, N.Y..

— (1968) Self-Reference and Self-Awareness, *The Journal of Philosophy*, Vol. 65, No. 19, pp. 555-567.

— (1970) Persons and Their Pasts, *American Philosophical Quarterly*, Vol. 7, No. 4, pp. 269-285.

SNOWDON P F 1995, 'Persons, Animals, and Bodies,' in J.L. Bermúdez, A. Marcel, N. Eilan (eds.) *The Body and the Self*, MIT Press, Cambridge (MA), pp. 71-86.

STRAWSON, G 2011, 'Hume's Appendix', in G Strawson *The Evident Connexion. Hume on Personal Identity*, Oxford, Oxford University Press.

STRAWSON, P.F. 1959, Individuals. An Essay in Descriptive Metaphysics. Routledge, London.

VENDLER, Z 1984, *The Matter of Minds*, Clarendon Press, Oxford.

WIGGINS, David (1976) Locke, Butler, and the Stream of Consciousness. And men as a Natural Kind, *Philosophy*, Vol. 51, No. 196, pp. 131-158.

WILLIAMS, B 1973a, 'Imagination and the Self', in B Williams, *Problems of the Self*, Cambridge University Press, Cambridge (UK), pp. 26-45.

— 1973b, 'The Self and the Future', in B Williams, *Problems of the Self*, Cambridge University Press, Cambridge (UK), pp. 46-63.

— 1973c, 'Are Persons Bodies?', in B Williams, *Problems of the Self*, Cambridge University Press, Cambridge (UK), pp. 64-81.

WITTGENSTEIN, L. 1958, Preliminary Studies for 'The Philosophical Investigations'. Generally Known as The Blue and Brown Books, Harper and Row, New York.

WOLLHEIM, R 1980, 'On Persons and Their Lives', in A. Oksenberg Rorty *Explaining Emotions*, University of California Press, Berkeley, pp. 299-321.